CREATIVE UNDERSTANDING

Philosophical Reflections on Physics

Roberto Torretti

Creative Understanding

CREATIVE UNDERSTANDING

Philosophical Reflections on Physics

Roberto Torretti

The University of Chicago Press
Chicago and London

Roberto Torretti is professor of philosophy at the University of Puerto Rico and editor of the journal *Diálogos*. Among his several books is *Relativity and Geometry*.

The University of Chicago Press, Chicago 60637
The University of Chicago Press, Ltd., London

Printed in the United States of America

99 98 97 96 95 94 93 92 91 90   5 4 3 2 1

Library of Congress Cataloging-in-Publication Data

Torretti, Roberto, 1930–
Creative understanding: philosophical reflections on physics / Roberto Torretti.
p. cm.
Includes bibliographical references (p.) and index.
ISBN 0–226–80834–3. — ISBN 0–228–80835–1 (pbk.)
1. Physics—Philosophy. 2. Physics—Methodology. I. Title.
QC6. T653 1990 90–11007
530´.01—dc20 CIP

♾ The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48.1984.

For Carla

ACKNOWLEDGMENTS

In writing this book I have been much assisted and encouraged by others, to all of whom I express here my warmest thanks.

I conceived the book and began working on it during a six-month visit to the Center for Philosophy of Science at the University of Pittsburgh in 1983-84. I am grateful to Larry Laudan and Nicholas Rescher, who, as directors of the Center, invited me to go there, and to Adolf and Thelma Grünbaum, who did much to make my stay intellectually fruitful and in every sense enjoyable. While in Pittsburgh I benefited from long conversations with David Malament, John Norton, and Richard Healey.

An earlier version of Chapter 1 was published in the *British Journal for the Philosophy of Science*, vol. 37, pp. 1–23 (1986). The editor, Professor G. M. K. Hunt, has kindly authorized me to print a modified version here. The first version of Chapter 3 appeared in *Diálogos*, Nº 48, pp. 183–212 (1986), and Nº 49, pp. 147–88 (1987), and I have permission to republish the portions that have not been rewritten.

John Stachel read the first draft of Chapters 1 and 2 and made some incisive comments which I hope I have duly taken into account. C. Ulises Moulines commented on the previously published text of Chapter 3 in two long letters which were most helpful to me when I rewrote Sections 3.3 and 3.4. Ronald Giere kindly clarified to me a point in one of his writings. Georg Fromm detected a mistake in Chapter 5. Jorge López instructed me about a basic fact of measure theory which I had not properly grasped. Two referees for the publisher pointed out quite a few errors which I have since corrected. They also made broad proposals for rewriting the book (in different ways) which, however, I did not venture to follow. Pamela Bruton edited the manuscript with unfaltering precision, greatly improving my use of relative pronouns and introducing some very welcome stylistic changes.

I typed the book myself, so I have no one to thank for that, but I am glad to record here my grateful admiration for the creators of the personal computer and the software that goes with it. I am especially grateful to Joe Carroll, for allowing me to use his equipment and teaching me how to print with it, and to Marc Cogan, for his prompt and friendly replies to my questions regarding the Greek font Kadmos.

I thank the staff of the University of Chicago Press for their pleasant and efficient handling of the book's publication.

The University of Puerto Rico granted me leave to go to Pittsburgh and has since repeatedly freed me from some of my teaching duties so that I could work on the book. I am, of course, deeply grateful for such continued support of my research; but above all I have to thank the University and the people of Puerto Rico for twenty years of complete academic freedom and almost complete tranquillity at a time when these prerequisites of scholarly work could no longer be taken for granted in my homeland.

My greatest debt is to Carla Cordua, who has long been to me a tireless source of light and strength. I am very happy to be able to dedicate this book to her.

CONTENTS

1 Observation

"Begin at the beginning!"—said the King of Hearts to the White Rabbit.[1] Many philosophers—call them *foundationists*—have been at pains to do so. They feared with good reason that if they let their thoughts take root right where they stood, in medias res, their quest for certainty would run afoul of a host of uncertified presuppositions. Reaching for the "unhypothesized First"[2] foundationists conjured up such fancy creatures as the bodyless ego and the unconnected simple sense datum and burdened their successors with thankless tasks, like proving the existence of the "external" world or "justifying" the way we ordinarily try to figure out the future from the past. I shall not go along with them. The understanding currently alive in language and other standard practices is the unfinished, unpredictably evolving outgrowth of partly random events, but we cannot do without it. We can work on it from within to broaden it and refine it, to fill in its gaps and to tie its loose ends, but there is no way that we can step outside it and build a better understanding from scratch. I must therefore rely on my own and the reader's command of some words and phrases and assume some truisms into which, so to speak, they fall of themselves. Of course, ordinary language—especially where it is not being used for survival—is not free of fog, and the line between incontestable truisms and disposable dogmas is often elusive. But the leeway that philosophy gains from such uncertainties should be husbanded with care. In particular, one ought not to cauterize the fuzzy edges of traditional meanings for the sake of clarity and precision, lest one wind up in yet another prim philosophical house of cards.

Science grows by observation. This commonplace I take for granted. For the foundationist who accepts it, its implications are plain enough: Observations must convey independent bits of knowledge, or else they would not deliver increments to science; hence, the systematic generalizations of science must be distilled from the aggregate of such particulars by some sort of logical alchemy; to establish its rules and to vindicate them is the main business of the philosophy of science. There is, however, a different and, to my mind, more promising philosophical approach to scientific observation: Just take it as it is currently practiced and understood, and try to tell from it what a science feeding on it can be. This approach can be traced back to Norwood Russell Hanson (1958). In the last few years it has been developed and applied in several excellent philosophical and historical studies.[3] In a similar spirit we shall now briefly review the main features of observation, to brace ourselves for the inquiries of the following chapters.

In Section 1.1 I bring to the fore two facets of ordinary observation, viz., physical interaction and awareness, and note that the former, but not the latter, is involved in every scientific observation. I introduce the expressions *personal* and *impersonal observation* to name observation with and without awareness and stress the role of general concepts in both forms of observation. Our conceptual grasp of physical objects will be the main subject of this book. With regard to it, foundationists traditionally raise two questions which I shall deliberately neglect: Where do the concepts come from? How can one be certain that a given concept is appropriate for a particular use? I comment briefly on these questions in Section 1.2. In Section 1.3 I consider another traditional question that we ought to come to terms with before proceeding with our inquiry: When personally observing a physical object, is one immediately aware of that object, or must one infer its presence and its features from directly perceived mental objects? Although this question cannot be satisfactorily dealt with in the space I can devote to it here, our short discussion should help to set, so to speak, the philosophical tone of the book and to get the reader attuned to it. The next three sections broadly examine our current understanding of observation as a physical process. Having made some very general considerations about it in Section 1.4, I compare personal with impersonal observation from this viewpoint in Section 1.5. Finally, in Section 1.6 I draw the conclusions I have been reaching for in this chapter: If every observation involves physical interaction between the observed object and a living or inanimate receiver, and every difference observed in the former must correspond to a difference recorded in the latter, then, obviously, a state of the receiver can supply information about the presence of a certain feature in the object only if—or insofar as—this feature is judged to be a necessary condition of that state. Consequently, the information we obtain by observation depends on our understanding of necessary connections in nature.

1.1 Two main forms of observation

Consider a few examples. A plainclotheswoman observes a slim clean-shaven young man come out of a house suspected of being a terrorist den. A nurse observes a mercury column rise and then slowly descend while she attends to the heartbeats she hears through a stethoscope, as she takes a patient's blood pressure. A tourist observes a small green caterpillar crawl slowly yet confidently on the piece of Danish pastry she has ordered for breakfast. In all these cases, a person, the *observer*, pays attention to something, the observed *object*—typically, a physical process or state of affairs, involving one or more things and their changing properties and relations—of which she is distinctly aware. Moreover, in all these cases the object physically interacts with the observer in a manner characteristic of the type of observation performed.

Some characteristic form of physical interaction is a necessary condition of the observation of physical objects, as it is currently understood. If the nurse in my example saw the mercury column fall with her eyes shut, one would say that she hallucinates it, not that she observes it.[4] Whether any physical interaction is required for the observation of one's own mind is a moot question which cannot be tackled here. However, to avoid cumbersome exceptions to what I shall be saying, I propose to use the word 'observation' in such a way that it does not cover the introspection of mental states and processes. (If the reader feels unhappy about this admittedly conventional restriction of the term's extension, he may add a constant prefix of his own choosing to all further occurrences of 'observation', 'observe', 'observable', etc., in this book.)

On the other hand, I should not say that every observation is made by a human observer who thereby becomes personally aware of the observed object. Indeed, most of the observations on which science thrives are carried out by artifacts designed to interact in a specific way with objects of some sort and to record the effects of the interaction so that human beings can eventually make use of them. The record may consist in a passing event (e.g., a click) or a lasting state (e.g., a pattern of black and white spots, as in an X-ray picture), of which a person can become aware by observing it in turn. But the record can also be inputted into a computer which combines many observation records, according to some preset program, in a computer output that a person can read. Now, I do not think that reading—unless it is *proof*reading—can be properly said to be a form of observation. You might observe a computer printout to see whether it is neatly printed or whether you ought to change the ribbon or make some other adjustment, but not to learn what the printout says. Consider, for instance, a scientist who peruses the output of a computer programmed to make weather forecasts from the data supplied by a widely distributed array of diverse meteorological instruments. Could one say that he is indirectly observing today's weather conditions? No more than one could say that he indirectly observes the weather conditions of July 11, 1953, when he rereads the record of the calculations he made himself with data gathered by personally reading the appropriate meteorological instruments on that day. And, of course, observation records can be used as feedback in an automated industrial or laboratory process without anybody taking notice of them. Evidently, many observations are now being performed in science and also in medicine, manufacturing, etc., that do not, and never will, involve observational awareness of their outcome, and hence should be described as observations without an observer.[5]

Let us therefore distinguish between *personal* and *impersonal* observation. In both forms of observation a result of a physical process is recorded in a distinct physical system I shall call the *receiver*. In personal observation, the receiver is the body—or should we say the nervous system?—of a human being who, upon reception of the record, becomes aware of the object of observation. In impersonal observation the receiver can be any of a wide variety of things—even a live human body on which, for instance, one tests a new drug or the tyrant's dinner. In all observations, the record of the physical interaction may last for years—as is the case with photographs—or just for a fraction of a second—as the neuronal excitation caused by a flash of lightning we see with our eyes—but it is never instantaneous. Even if, like everything else in nature, the physical state constitutive of an observation record is continually changing and is thus never quite the same, it is only by virtue of its more or less lasting sameness—as far as it goes, and however it may have to be understood—that it is the record of just that observation. Instantaneous states and events can no doubt be detected and measured, but only through more or less stable receiver records.

The awareness involved in personal observation is a peculiar mode of consciousness that differs, say, from fear or recollection, but also from other forms of perception, such as watching a movie or basking in the sun on a beach. Like other deliberate, attentive modes of consciousness, observational awareness is *self*-conscious: the observer must be aware of observing, or else he cannot properly be said to observe. He may, indeed, be so absorbed by his task that he becomes oblivious of himself; but even then he will be aware of the observed object as something that is being *observed*—not, say, imagined or merely thought about. Observational awareness is always, of course, awareness *of* something, the object of the observation, a complex of changing and unchanging features singled out from life's flux. Characteristically, the observer grasps the object as a particular instance of a universal. I do not mean to say that it is impossible to pay attention to an individual as such, and not as a member of a class. When talking to a close friend, or laughing together, or holding her in my arms, I am often aware only of the unique individual who is with me, and not of any universals that she instantiates. Awareness of individuality in its irreplaceable uniqueness is, I should say, a necessary ingredient of any genuine personal relationship. But such awareness is not observational, and as soon as one begins to observe one's partner, her teeth, her accent, her syntax, one does in effect subsume her under a general concept. Thus alone can the observer establish what is being observed. As the observation proceeds, the concept is further specified and articulated and may also be revised or replaced. For, far from being incorrigible as it has sometimes been suggested, a personal observation undergoes continual revision while it lasts.

In impersonal observation general concepts function in a rather different, but no less decisive, manner. Conceiving the object is not here an integral part of the act of observation. But in order to constitute an *observation,* a physical process must be conceived as such by the people who set it up or who intend to use its results. Recorded receiver states cannot disclose anything about a purported object of observation unless this object and the receiver are appropriately distinguished and the interaction between them is somehow understood. We shall have more to say on this later. For the time being, I conclude that, since concepts go into every observation, empirical knowledge is intellectual through and through. Kant said it bluntly: sense awareness without concepts is blind.

1.2 Conceptual grasp of the objects of observation

In the remaining chapters of this book we shall in one way or another be dealing with the conceptual grasp of physical objects, especially as it is practiced in mathematical physics. Here I only wish to mention—and to dismiss—two problems regarding that grasp which have greatly exercised philosophers of the foundationist persuasion.

One is the problem of noogony,[6] or the origin of concepts. They cannot *all* be obtained *from* observation by the standard procedures of comparison, reflection, and abstraction. In each observation some concepts must be at work from the very outset. Each time a judgment is revised, some concepts must remain stable. Thus, every exercise of our understanding involves concepts which are, one may say, locally a priori—i.e., presupposed and taken for granted for the occasion. Clearly, then, the doctrine that all concepts proceed from observation would entangle us in an infinite regress. This does not imply that any concepts are permanent, or global in scope. But some of them, at least, must be born ex nihilo. Whether we lay them at the door of Mother Nature or of the Muse of inquiry or of God Himself is presumably a matter of taste. Whatever one's choice may be, it appears that in dealing with noogony one cannot avoid mythology.

The second problem concerns the appropriateness of any given concept for a particular use. Every conceptual grasp of an object of observation is liable to revision and correction in the light of other observations; indeed, as I noted above, *self*-correction is the very soul of personal observation. What justifies our preference for one concept over another? How can we judge that we have achieved a better conceptual grasp? Foundationists miraculously solve—or rather dissolve—this problem by appeal to the dogma of Immaculate Perception. According to it, we observe virgin data, unpolluted by our deceptive intelligence, and can adjust our concepts and judgments to them. Now, if one could ever make a perfectly self-contained observation, not signifying anything beyond itself, there would certainly be no means—and no motive—for revising it. One would just blow away the conceptual chaff and leave the observational grain alone in its splendid isolation—and irrelevance. But in fact no observation is thus self-contained.[7] Each one of them is constitutively linked by concepts to other observations welded into a complex network of assumptions and beliefs, together with which it gives rise to a wealth of expectations.

Failure of expectations is perhaps, in the end, the main inducement for revising and correcting our observations. But earlier observations can assist in the correction—and even lead to the rejection—of subsequent ones, if they are more detailed or more careful or more consonant with one another. Consonance and detail furnish unquestionable, more or less unambiguous criteria of preference. But to say that an observation is more careful than another one would seem to presuppose the very choice that we seek to vindicate. However, some observational procedures may well be deemed more careful than others if they normally lead to more successful expectations. Moreover, our well-corroborated understanding of the physical processes of observation provides definite and, for that same understanding, impeccable grounds for assessing the reliability of observations. How one goes about using such diverse criteria in the progress of experience is well known from our daily lives. More sophisticated examples are provided by the history and current practice of scientific research. Beyond such bland generalities, philosophy has very little to say about the scientific procedures for collating observation data and the criteria for judging their worth. They pertain to the methodology of each field of inquiry and are decided upon by the practicing experts in the light of their current understanding of the matter at hand.

One philosophical generality should never be forgotten: if all our knowledge of physical objects is corrigible, it must be self-correcting, for there is no outside authority to which one could turn for help. Quine's famous dictum that "our statements about the external world face the tribunal of sense experience not individually, but only as a corporate body" (Quine 1961, p. 41) is apt to be misleading. For in the trial of empirical knowledge the defendants are at once the prosecution, the witnesses, and the jury, who must find the guilty among themselves with no more evidence than they can all jointly put together. This truism can be stated less aggressively, and perhaps more sensibly, as follows. In philosophy, things are said to be as we understand them to be, but we are well aware that they might not be that way. Such awareness, however, does not result from our transcending our understanding and glimpsing at things beyond it. It simply expresses our discontent with our own views and thoughts, which we feel to be incomplete, murky, or plainly inconsistent. But improvement can only be had by thinking harder, and we alone must see to that. (On the other hand, should one ever achieve perfect intellectual self-satisfaction, one would find no occasion for distinguishing between truth and appearance, between the way things are and the way one thinks and says they are—except, indeed, to contrast one's present knowledge with one's former ignorance and with the beliefs of others.)

1.3 On the manifest qualities of things

In personal observation, are we observationally aware of the object of our attention or merely of its effect on our minds? This is a question I think we ought to face before going any further. The answer will not change the way physics is done and may not be required for its proper elucidation, but a discussion of the question itself, its motivations and presuppositions, and a brief but clear statement of my own stance on the matter might prevent misunderstandings of what I shall be saying later.

To someone untainted by philosophy the question sounds silly, a typical example of the idle sport practiced by tenured professors on their captive audience. Isn't the juicy pineapple chunk I am now chewing the very same thing that feels cold and tastes sweet on my tongue, indeed the same one that looked yellow and smelled of pineapple a moment ago, when I carried it with a fork to my mouth, in front of my eyes and my nose? Normally, I should not have the slightest doubt about this, if 'is' and 'same' and the other ordinary English words employed are being used properly. It could happen, of course, that I fell into a swoon while still holding the pineapple in my hand and that another piece was put into my mouth while I lay unconscious. But such a situation can only be contemplated and diagnosed by contrasting it with the more common one, in which I eat what I pick and I do not pick anything without first seeing it. Of course, the food must act on my senses, or I would not be aware of it. But it is *the food* that I am aware of, as it stands alluringly on the plate, and later, as it yields to my teeth; not the changes caused by it in the cones and rods in my eyes and the papillae on my tongue, or in this or that lobe of my brain, or—as some are fain to say—in my soul.

Yet a venerable philosophical tradition maintains that, among the apparent properties of a body, only size and shape, position and speed can really belong to it and that the colors, sounds, tastes, odors, etc., which it sports are only the effects that its action on our bodies causes on our minds. The oldest known statement of this thesis was made by Democritus of Abdera (ca. 375 B.C.):

> νόμῳ γλυκὺ, νόμῳ πικρόν, νόμῳ θερμόν, νόμῳ ψυχρόν, νόμῳ χροιή, ἐτεῇ δὲ ἄτομα καὶ κενόν.
>
> By custom, sweet; by custom, bitter. By custom, hot; by custom, cold. By custom, color. In truth: atoms and void.
>
> (Democritus, in Diels-Kranz, 68.B.9)[8]

Democritus' motivation is well known. He had learned from Parmenides that being cannot come from not-being or not-being from being, but he would not stomach Parmenides' denial of variety and change. So he made allowance for not-being (μηδέν) in the guise of the void (κενόν), to make room for his atoms—tiny unborn unchangeable indestructible Parmenidean beings—to differ from each other and to move. But the face which this reality presents to us he still regarded as a man-dependent appearance, to be accounted for by the changing configuration of the atoms that make up our souls.

Democritus' curt dismissal of the manifest qualities of bodies found little following among latter-day atomists in antiquity,[9] but was revived and collectively embraced by the founders of modern science. P. M. S. Hacker (1987, Chapter 1) has gathered a series of passages from Galileo, Descartes, Boyle, Newton, and Locke. Allow me to quote a few of them for the benefit of readers who do not have Hacker's book at hand.[10] The following is from *The Assayer* (1623):

> Che ne' corpi esterni, per eccitare in noi i sapori, gli odori e i suoni, si richiegga altro che grandezze, figure, moltitudini e movimenti tardi o veloci, io non lo credo; e stimo che, tolti via gli orecchi le lingue e i nasi, restino bene le figure i numeri e i moti, ma non già gli odori né i sapori né i suoni, li quali fuor dell'animal vivente non

> credo che sieno altro che nomi, come a punto altro che nome non è il solletico e la titillazione, rimosse l'ascelle e la pelle intorno al naso.
>
> I do not believe that anything is required in external bodies besides their size, shape, multitude, and motions, fast or slow, in order to excite in us tastes, odors, and sounds; and I think that if ears, tongues, and noses are removed, the shapes and numbers and motions will remain, but not the tastes nor the odors nor the sounds. Apart from the living animal, the latter—I believe—are nothing but names, just as tickling and titillation are mere names if the armpit and the skin lining the nose are removed.
>
> (Galileo, *Il Saggiatore*, §48 [EN, VI, 350])[11]

Robert Boyle made the same claim in *The Origin of Forms and Qualities according to the Corpuscular Philosophy* (1666):

> We have been from our infancy apt to imagine, that these sensible qualities are real beings in the objects they denominate, and have the faculty or power to work such and such things [. . .] whereas [. . .] there is in the body to which the sensible qualities are attributed, nothing of real and physical, but the size, shape and motion or rest of its component particles, together with that texture of the whole, which results from their being so contrived as they are; nor is it necessary they should have in them anything more, like to the ideas they occasion in us.
>
> (Boyle, WW, II, 466)

He admits, however, that the ordinary names of colors, tastes, etc., may be used "metonymically" (Boyle, WW, II, 7) to designate those textural features by virtue of which a body can effect the homonymous sensible qualities in our minds.

> I do not deny but that bodies may be said, in a very favorable sense, to have those qualities we call sensible, though there were no animals in the world; for a body in that case may differ from those, which now are quite devoid of quality, in its having such a disposition of its constituent corpuscles that in case it were duly applied to the sensory of an animal, it would produce such a sensible quality, which a body of another texture would not: as though, if there were no animals, there would be no such thing as pain, yet a pin may, upon the account of its figure, be fitted to cause pain in case it were moved against a man's finger.
>
> (Boyle, WW, II, 467)

In this, Boyle was generally followed by later writers, notably Locke, who liberally granted us the right to go on saying that grass is green and ice is cold because of the standing disposition of such things to make us see the color green or feel a sensation of cold. Boyle and his followers did not explain why pins are never said to be painful, let alone pained, in that metonymical sense in which Henri Rousseau's paintings are colorful and colored. Presumably they counted this among the vagaries of standard English.

The philosopher-scientists of the 17th century had no truck with Parmenides, and their position concerning manifest qualities was chosen on epistemic grounds. They had this vision that the Book of Nature is written in the language of mathematics (Galileo, *Il Saggiatore,* §6 [EN, VI, 232]), but mathematics as they knew it could deal only with the size, shape, position and change of position of bodies. Modern differential geometry has taught us to conceive of mass and force and temperature and heat flow as "geometric objects" on a differentiable manifold, but such notions were still a long way off at that time. So if the new mathematical science of nature was to tell us how things really are, things had first to be stripped of their manifest qualities. Descartes expressed it with his usual lucidity: Material things can be clearly and distinctly conceived, and thus their nature known with certainty, only "insofar as they are the subject-matter of pure mathematics" ("quatenus sunt puræ Matheseos objectum").[12]

But, of course, if the visual redness of a red cube is just a mental effect of its presence in our line of vision, its visual shape or position will not be otherwise. One may concede, perhaps, that visual shapes and positions resemble—whereas perceived colors do not resemble—their homonymous counterparts in the body we see. But they cannot themselves belong to it, unless the colors do. For such visual shapes and positions are identical with, or are compounded from, the shapes and positions of visual displays of color.

As analogous considerations are extended to all the manifest qualities of things, bodies fade into inferred entities, whose real presence—endowed with such-and-such properties and relations—is postulated in order to account for the perceived features of the mental kaleidoscope of sense appearances, which are all that, properly speaking, we see and hear and taste and smell. In his recent defense of such a view, Frank Jackson assimilates the philosopher's postulation of bodies as the source of mental "sense-data" to the physicist's postulation of molecules to account for the ostensible behavior of gases.[13] There is, however, a great difference between them. If you postulate the existence of small bodies as components of a larger one you observe, you are making a hypothesis about its parts, which must exist if it does. But if you maintain that all you are observationally aware of are mental objects, and postulate bodies as causes for them, you are introducing an entirely new category of being for which, by your own claim, you have no proper evidence.

I shall now describe an imaginary—though practically unimaginable—condition in which an observer might plausibly understand himself as being observationally aware only of his own mental states. The difference between that condition and our own will make clear, I hope, why we cannot persist in speaking, for any significant length of time, as if our senses made us aware only of so-called sense appearances. Consider a purely contemplative observer who sees static scenes, one after the other. He would have little or no inducement to analyze the scenes into parts or to associate parts of different scenes unless such parts were equal; and even if they happened to be so, he would have no reason for distinguishing the object of observation from its momentarily perceived aspects. Suppose now that the scenes observed change gradually and flow into each other, as in a motion picture. The observer could then perhaps discern patterns in the flow and come to view parts of successive scenes as diverse aspects of the same object. Such an object, however, would be no more than the series of its presentations, or rather the law of that series. A Humean analysis would unmask such laws, exposing them as mere habits. We can add sound and even smells to the motion picture without essentially changing the situation. We humans differ, of course, from a purely contemplative observer in that we have an interest, often a vital one, in the objects we perceive and are sometimes able to change them. But even if we let our fictitious observer resemble us in this, if we allow, say, some of the movie sequences, which are all he is aware of, to be pleasant or painful, and if we let him will and occasionally achieve the removal of pain, the renewal of pleasure, he still would not be one of us. He lacks the complex array of muscular, postural, thermal, tactile, or—to name them all by a single Greek word—*haptic* experiences in which we perceive ourselves as bodies incessantly interacting with other bodies, dangerously exposed to them, and also, through that very interaction, capable of manipulating them and observing them. The pencil I hold in my hand and press between my fingers, the chair I sit in, the table I write on, are grasped in observation as true bodies because through the pressure I exert on them, the movements I make against them, the thermal gradients they generate on my skin, I sense their bodily presence on a par with my own. Dr. Johnson refuted Berkeley by kicking a stone. The Greeks fought Pyrrho's scepticism by letting a dog loose on him. Professors smile with condescension on such wordless arguments, but there is a wisdom in them. Macbeth would clutch the dagger that he saw before him or else dismiss it as "a dagger of the mind."

Awareness of our interaction with the bodies surrounding us is the key to our construal of personal observation as a physical process, with our body as the receiver. It is convenient to recall how this construal introduces a measure of order and consistency into the diverse and often baffling appearance of things. If the observer becomes aware of the physical objects about him by their action on his body, his observational awareness must depend not only on the objects themselves but also on the condition of his body and all other circumstances influencing the observation process. Thus, our grasp of the physical basis of vision enables us to understand why a Gothic steeple should look different through the fog and under a blazing sun, why a pencil should show a kink when partially submerged in a glass of water, why the police van catching up with my car from behind should turn up in the mirror in front of me, why a supernova should now flare up in the sky in the direction where it faded out forever several million years ago. On similar grounds we can account also for our seeing visibles (and hearing audibles, etc.) that are not judged to be an aspect of anything, such as the red, semitransparent disks that we see wherever we direct our eyes after we have been looking intently for a while at a strong source of light, or the colorless little worms that we see wriggling about in the air if we stare at a bright cloudy sky. Since visual (acoustic, etc.) awareness closely depends on the state of the body, it is to be expected that it will often be stirred by changes in that state which are not a part of any process of observation (just as, say, a short-circuited loudspeaker will emit a noise which is not a part of any music being played). It is fortunate, indeed, that such occurrences, though frequent, rarely become obtrusive. But it is a perversion of philosophy to choose such marginal events as the prototype of all our sense experience and then to wonder how it may come to pass that by far the greater part of it is so neatly ordered as a display of physical objects. In fact, outside this order in which we normally perceive things in their manifold aspects, it is hard to conceive that there could even exist an awareness of objectless *sensibilia.*

Philosophers sometimes run into difficulties with the manifest qualities of things because, obfuscated by half-unconscious *theologoumena,* they unwittingly set standards of determinacy for things which the latter in effect do not meet. They persist in understanding things as *res*—in the scholastic medieval sense—or things-in-themselves, when all one ever meets and has to do with are *pragmata,* or things-in-our-environment. What one ordinarily means by a *thing* is, of course, colored or transparent, quiet or noisy, tasty or insipid, pretty or ugly, and though modern physics has thrived by methodically neglecting such features, its mathematical constructs are designed to represent and to assist us in understanding and handling the very things that sport them.[14]

1.4 Our understanding of the process of observation

In personal observation the observer apprehends his own body in physical interaction with the objects observed. Observational awareness never lacks this feature, at least where haptic perceptions are at play. Throughout our lives this is practically always the case, so it is no wonder that, in ordinary usage, the statement that a person *x* observes a thing or event *y* implies the statement that *y* causes *x* to be in a state in which she succeeds in observing it. Indeed this usage extends to all modes of observation, visual, auditive, etc., even where no reference is made to haptic awareness. I do not take this linguistic practice to mean that, say, purely visual observations—if such exist—must involve a claim to being caused by their objects (except perhaps when they are on the verge of being painful due to excess of light, in which case vision becomes proprioceptive like touch and kinesthesia).[15] But visual observations are made by us, men and women of flesh and blood, who must sit or stand or walk or run or turn or stoop or stretch or, at the very least, strain our eyes to see. Haptic awareness is thus pervasive and discloses, in one way or another, that we are committed to the physical world. Our everyday handling—holding, pressing, pulling, pushing, twisting—of all sorts of bodies and our continual exposure to bumping and falling, heat and cold, wind and water, light and noise, furnish the prototypes of our original notions of physical existence and physical action. It is therefore most unlikely that we shall ever find occasion of rejecting our grasp of ourselves as bodies interacting with other bodies. Philosophical attempts at replacing this ingredient of our self-understanding have hitherto been little else than exercises in the abuse of language. As John Dewey wrote at the beginning of his *Logic*: "It is obvious without argument that when men inquire they employ their eyes and ears, their hands and their brain."[16]

Yet while men have never seriously hesitated in their grasp of observation as a physical process, their general understanding of such processes has undergone great changes. For example, Aristotle conceived of a manner of physical action that was designed to account for perception and observation. By virtue of it, the constitutive "form" of the observed object could be transmitted "without matter," through an appropriate intervening medium, to the "sensitive faculty" of the observer.[17] This doctrine was taught at school to the founders of modern science, who later rejected it and replaced it with a different conception of physical action which in part revived pre-Aristotelian notions. Towards the end of the 17th century the new conception had taken such hold of the best minds in Europe that, for example, John Locke "found it impossible to conceive that body should operate on *what it does not* *touch* [. . .], or when it does touch, operate any other way than by motion."[18] Whence, when he comes to consider "how bodies produce ideas in us," he declares that it "is manifestly by impulse, the only way which we can conceive bodies to operate in."[19] This early modern idea of physical action was considerably modified by successive generations of natural philosophers, first by the 18th century theorists of instantaneous action at a distance, then in the 19th century by the creators of field theory. One capital ingredient of it survives, however, to this day: for us, as for Descartes, Huygens, etc., all physical action boils down to a transfer of momentum—or, as we would rather put it now, of four-momentum.[20]

The modern philosophy of nature has presided over great advances in the physiology of perception. It has also been associated, from its inception, with the modern development of means and methods of impersonal observation, which not only have tremendously expanded the scope of our knowledge but should also help us, through our growing familiarity with them, to achieve a better grasp of the nature of personal observation. On the other hand, the modern idea of physical action has burdened us—also from its inception—with the so-called mind-body problem. For, as the 17th century occasionalists were quick to see, transfer of momentum will neither account for nor be explained by a change of mind. All the attention devoted to the problem since Descartes has not brought us any nearer to understanding how a man's decision can initiate a definite outward flow of energy and momentum across his skin, or how an inward energy-momentum flow across it can modify his state of awareness. And we still do not know how to coordinate our particular states of awareness of observed objects with any well-defined, particular effects of the action of such objects on our bodies. It is unlikely that this rift between the two sides of observation can be closed without some radical, incalculable innovations in our understanding of physical action. But since our current understanding lies at the heart of so much valuable knowledge, there is little inducement to change it.

Even if our present understanding of the observation process is thus limited and beset with difficulties, we are deeply committed to it, and we cannot well imagine how some of its implications could be denied. Thus it seems clear that, no matter how we conceive physical action, in every observation the observed object interacts with a receiver. Such interaction is critical to the acquisition of knowledge by observation, for the observer cannot ascertain any more features of the observed object than become discernible to him through their recorded effect on the receiver. Indeed, a state of the receiver can furnish information about a feature of the object observed only to the extent—and within the range of ambiguity and imprecision—that the said feature is, under the circumstances, a necessary condition for the attainment of that state. The receiver's "power of resolution," its capacity to separate—or its tendency to blur—the imprint of different attributes and states of the object, is a measure of its cognitive value. From this point of view, impersonal observation, carried out by means of an increasingly diverse and efficient panoply of precision instruments, enjoys a distinct advantage over personal observation.

1.5. Personal versus impersonal observation

Observation processes have their peculiarities, without which they would not serve their purpose, but they are not generically different from other physical processes which are not observational. Observational interaction instantiates the same types and obeys the same laws as ordinary physical interaction. Indeed, the development of impersonal observation in the modern age could only get under way on the understanding that such was the case. Observation devices exploit known properties of well-typified natural processes for the sake of collecting information. Inference from the state of the receiver to the state of the observed object must rest on our knowledge of those properties, and can therefore hold good only if observation processes are not, physically speaking, a class apart.

Nonetheless, observation processes do differ from their nonobservational analogues in that they are ordered to an end: they are always embeddable in a quest for information. It is a requisite of this teleological order that, among the many factors that contribute to a physical process of observation, some should stand out as the objects of observation and their observed features, while others constitute the receiver and its data-recording states.

In impersonal observation, the receiver is usually artificial and is singled out by its human manufacturer. It is expressly designed to register the interesting effects of the intended object of observation, which has been previously singled out by some human research project. Since the object-receiver interaction is nevertheless immersed in nature's flux, great ingenuity must usually be devoted to filtering out the "noise" that hinders the clean flow of information from the object to the receiver. The status of these several items is indeed notional and depends on the project which the observation is meant to serve. Thus, by timing the eclipses of one of Jupiter's moons—intermittently hidden behind the planet—you can ascertain its period and thence, by Kepler's Third Law, its average distance from the planet, provided that you know the speed of light and use it to correct the anomaly of the observed period due to the Doppler effect consequent upon the relative motion of Jupiter and the Earth. If, on the other hand, you note the said anomaly but do not know the speed of light, you can, like Ole Rømer, use the timing of the eclipses to measure it—provided that you know the relative velocity of Jupiter and the Earth.[21]

The significance of the receiver's states is a matter of interpretation, depending, of course, on the circumstances of the observation—a thermometer reading will not tell us much about a child's fever if, on coming out of the child's mouth, the thermometer has fallen into a bowl of hot soup—but also, decisively, on the observer's understanding of the experimental situation. On the frontier of research, such understanding is apt to be flimsy. Thus, for example, the negative result of Michelson's famous attempt to measure the relative motion of the Earth and the ether was understood to indicate (*a*) that the ether is dragged by the Earth's atmosphere, the laboratory walls, the protective box in which Michelson's apparatus was enclosed, etc. (this was Michelson's own conclusion in 1881); (*b*) that the motion of the apparatus across the ether modifies the molecular forces that hold its parts together, shortening one of its perpendicular beams while merely narrowing the other (this was independently suggested by Fitzgerald and Lorentz); and (*c*) that we live in a Minkowski spacetime in which light pulses in vacuo follow null worldlines, so that the speed of light measured in an inertial laboratory in which time is defined by Einstein's method is the same in every direction. Or, to mention another, more recent example of an observation with positive results: the isotropic background noise that has been recorded in ultrasensitive microwave radio receivers around the world since 1965 and is generally regarded as the effect of thermal radiation of approximately 3 K, is understood as the manifestation of a very hot early global state of the universe. However, this cosmological reading of the phenomenon would have to be dismissed if it were determined that in another galaxy the noise is absent or is significantly anisotropic (or if it had turned out that outside the Earth's atmosphere its intensity does not peak at the frequency prescribed by Planck's law of thermal radiation).[22]

While the informative aim of an impersonal observation accrues to its underlying physical processes by human initiative, such a goal is, so to speak, endogenous to personal observation. Here the receiver has not been segregated from the mainstream of nature for fact-gathering purposes by an external agency but has grown of itself into a distinct, fairly stable physical system, suitably disposed to pick out specific effects of its interaction with specific objects. The information-bearing receiver states are not presented on a dial to the observer's interpretative acumen but translate spontaneously into observational awareness. The objects of personal observation do not have to be inferred from the states they induce in the receiver, for they are simply and straightforwardly perceived. In fact, it is rather from his direct awareness of them that the observer eventually learns—by inference—what receiver states are instrumental to their observation. Thus we have come to know that—though we are still quite incapable of explaining how—the recorded difference of less than 1/3,000 s between a sound's arrival in our left and in our right ear enables us to distinguish the direction from which the sound came; that our visual awareness of the volume of nearby bodies rests on the slight difference in the optical input from such bodies into each one of our eyes; that our sense of balance and orientation in the gravitational field in which we live depends on the flow of liquid along the sensitive walls of the semicircular canals in the internal ear. In inquiries leading to these and other results about the material conditions of perception, the physical objects of our perceptual awareness are the grounds, not the goals, of inference. Indeed the very notions of physical object, physical state, physical process—sophisticated though they have grown through the exertions of modern scientific thought—are rooted in the manner in which men and women, physically interacting with their surroundings, naïvely articulate their awareness of that interaction.

We normally have a more or less definite grasp of the objects of our personal observations and of their relations of place and time, and in some cases also of their causal relations with our bodies. This grasp is the source from which the theory and practice of impersonal observation ultimately draw their sustenance and motivation. Thus, personal observation may justly claim metaphysical priority over impersonal observation. But that does not bestow on it an epistemic privilege with respect to the latter. For personal observations and the "natural," unreflecting grasp of things that goes with them are both fallible and corrigible and are being continually rectified and qualified, not only by mutual comparison but also in the light of impersonal observations. Thus, we habitually compare the readings of outdoor thermometers or of wristwatches with the estimates of air temperature or of time elapsed based on our feelings—a practice which not only serves to control and to correct such estimates, but can also contribute to improving their accuracy. Personal observation is not only not superior to impersonal observation as a source of knowledge about physical objects, but, in both scope and precision, it is on the whole markedly inferior. The confusion that still prevails in some philosophical circles on this fairly obvious matter is due perhaps to a vicious craving for certainty. Of course, such craving will never be satisfied by impersonal observation, with its intricate scaffolding of theories. But neither can it be quenched by contracting one's knowledge claims to the bare subsistence level of commonsense judgments and naked eye observations.

Human perception must indeed always intervene at some stage of the harvest of impersonal observation data for use in science. Should not this obliterate the superior precision and reliability of those data? After all, a system for the transmission of information cannot perform better than its weakest link. However, the human sensors are not equally deficient at every task. They are rather bad for discriminating weights or temperatures or light intensities, and they are utterly useless for detecting small changes in atmospheric pressure; but they are pretty good for apprehending neatly printed digits and may be trusted to note a coincidence between a pointer and a thin black line on a white dial. Observation devices are designed to translate the often imperceptible effects of the observed object on the receiver into such easily perceivable receiver states. That persons should thus learn through their senses the outcome of impersonal observations has led some philosophers to think that a faithful description in plain everyday language of the apposite sense experiences can give the full "cognitive meaning" of the statements, couched in esoteric, "theoretical" terms, in which scientists normally report their findings. Of course, in real life things stand just the other way around: digital and pointer readings get their distinctive interpretation from the theory of the respective instruments, and without it they all look quite insignificant and very much the same.

1.6 On the relation between observed objects and receiver states

No difference can be observed in an object that is not recorded as a difference in the receiver. This principle is central to our current understanding of observation;[23] and it does not seem possible to deny it, no matter how we revise or refine that understanding. Indeed, the principle is so deeply ingrained in our language that we would never be said to *observe* a change we know to occur in the object, but which our bodies and the instruments at our disposal do not reflect.

It follows that in any personal observation receiver states must mediate between the observed features of the object and the observer's perception of them. We are far from understanding the relation between those states, of which we are mostly unaware, and our awareness of the objective situations they disclose. That there is no simple correspondence between the information-bearing states of our sense organs and any relevant states of the mind can be readily gathered from the examples of stereophonic and stereoscopic perception mentioned in Section 1.5. Only by sinking the cognitively significant receiver states deeper and deeper into the unexplored recesses of the brain can one hope to map them one-to-one onto the contents of our sense awareness. As neurology advances, such terræ incognitæ become increasingly unavailable, and one sees ever more clearly that a mind-brain isomorphism, if at all possible, can be established only on the basis of a thoroughly innovative, physically unorthodox description of the brain. On the other hand, the relation between the said receiver states and the matching features of the object can be handled by the standard methods of physics. In this respect there is no essential difference between personal and impersonal observation. And indeed practically all progress in the physiology of perception, since Kepler first conceived the eye on the analogy of the camera obscura, has been achieved by treating the organs of sense as impersonal receivers.

Object-receiver relations in personal and impersonal observations take varied forms and their study pertains to diverse fields of science. But they all share at least one common trait. *A receiver state conveys information about the presence of a certain feature in an object only if—or insofar as—this feature is judged to be a necessary condition of that state.*[24]

Consider impersonal observation. Although there may be no difficulty in classifying and recognizing observationally significant receiver states, a definite receiver state often does not unambiguously point to an equally definite feature in the object. That state may normally arise due to several conditions, some of which may not even involve the intended object of observation. (Precision measurements can be severely impaired by thermal variations in the instruments employed.) But even where such perturbing factors are negligible, the distinguishable states of the receiver may not suffice to discriminate between significantly different properties of the object. A gray shadow on a medical X-ray picture can reflect all sorts of conditions in the patient's body. To judge what is actually disclosed by it, an observer must rely on his experience of similar X-ray pictures and on his general knowledge of medicine. A coupled pair of spots in a telescope photograph of a piece of sky is usually taken as evidence that in the direction of those spots there are two, possibly associated, astronomical light sources; but the spots might exceptionally be caused by a single source, if the beam of light it sends towards us is split, on the way to our telescope, by a gravitational lens. To decide that the latter is indeed the case, a scientist must examine the circumstances in the light of gravitational theory.

There are, indeed, plenty of cases in which the record of an impersonal observation tells a person exactly what she wants to know about an object, although she has no inkling of how the observation works and of what precisely is recorded by the receiver. Thus, if, blindly following the instructions in a manual, I connect the terminals of a voltmeter to the knobs on the upper face of my car's battery, the position reached by the needle on the voltmeter's dial will let me know without further ado whether the battery is strong enough to start the car promptly on a cold morning. I require no theory, almost no experience, and very little judgment to draw the appropriate inference from the actual reading. Most of us ordinarily employ instruments of observation to learn about our surroundings in such a thoughtless way. But we can do so only because a vast repertoire of object-receiver correlations has been firmly established by scientific and technological research. Such research is all but thoughtless. It does not simply proceed by trying out any old instrument on a class of objects and setting up by straight-rule induction[25] a correspondence between the alternative states of the former and the interesting differences among the latter. The impersonal receivers in current use in all walks of life have for the most part been painstakingly developed in the light of scientific theories which entail certain necessary connections between diverse features of interest in our environment and directly observable receiver states. In Chapter 5 we shall consider what type of necessity scientific theorizing discovers—or should we say induces?—in nature. But it should already be clear that impersonal observation is impossible without it. A particular receiver state can disclose a particular state of affairs only if the latter is, under the circumstances, a necessary condition of the former. To judge it so, one must grasp them both as instances of general types which stand to one another in suitable relations of entailment. Such typifications are not ready-made but are the product of scientific thought. We may indeed unreflectingly profit from the impersonal observations with well-established significance which are taking place all about us. But we could not without reflection and theory-guided invention have brought them under way.[26]

So-called accidental discoveries might seem to be exceptions to this rule but ultimately tend to confirm it. Thus, for example, the first observation of radioactivity was recorded in a photographic plate stored with a preparation of uranium salts from 27 February to 1 March 1896 inside a drawer in Henri Becquerel's laboratory. The plate was exposed notwithstanding the absence of light in the drawer; but it took Becquerel's alertness and preparedness—he himself had mounted the uranium salts on the plate to study their phosphorescence under sunlight and had stored them in the drawer while waiting for propitious weather—to grasp as an observation record what another one would have discarded as a spoiled plate.

Physically, personal observation is no different from impersonal observation. A person cannot become aware, by observation, of a change in an object unless the latter effects a change in her body. A state of a human body cannot convey information about a feature of its surroundings unless this feature is, in the circumstances, a necessary condition of that state.[27] However, not every state of the body is a source of observational awareness; nor do those that are disclose every one of their necessary conditions. Observational awareness is selective: the observer's attention, guided by his interests and preconceptions, falls at any given time only on a small part of the current range of his consciousness. Observational awareness is self-transcending: it is no mere epiphany of organic states but the grasp of an object against the background of a world. Hence, while in impersonal observation the facts of the matter must be inferred from a suitable description of the receiver states in the light of scientific theories and a general assessment of the circumstances (or by means of the "inference tickets" supplied by the user's manual that comes with the instrument), in personal observation the actual presence of such-and-such an object is not a conclusion to be drawn deductively or inductively from the momentary state of one's body, for we are, so to speak, preprogrammed to jump to it straight away. (See Fodor 1984.) The observer's grasp of the object can be rectified to comply with earlier or further experiences, with scientific theories, or even with philosophical criticism. But it cannot be suppressed from observational awareness without destroying the latter's observational character. Thanks to this grasp of the environment in which his body is placed, the human observer develops an understanding of observation as a physical process and devises increasingly sophisticated theories about object-receiver links. Such theories are not required to get personal observation going—indeed, they would not even be possible if observational awareness did not precede them—but they are certainly apt to modify our grasp of what we observe personally.[28]

2. Concepts

In this chapter we begin our exploration of creative understanding in physics. In Section 2.1 I take a new look at the familiar view of scientific explanation as inference. I contend that such explanations require a rethinking of the facts, in order to bring them under the scientific theories that explain them. In Section 2.2 I illustrate this with some examples from Newton. Section 2.3 raises a question we must face if the facts of observation are grasped and regrasped under changing concepts: "Can the facts remain the same as the framework of description varies?" A negative answer to this question would warrant the so-called incommensurability of scientific theories, proclaimed in the 1960s by T. S. Kuhn. Section 2.4 criticizes two classical ways of forestalling such incommensurability, favored, respectively, by Kant and by Carnap. Section 2.5 studies the internal connection between two incompatible theories, one of which arises through criticism of the other. Section 2.6 discusses the theory of meaning introduced (and subsequently abandoned) by Hilary Putnam to rescue the stability of reference under radical conceptual change. Section 2.7 elucidates the notion of a conceptual scheme, implicit in the problem of incommensurability, and proposes a new approach to it which should go a long way to solving that problem. The mathematical appendix in Section 2.8 sketches the notion of structure that will pervade Chapter 3 but that is already employed in the present chapter for "speaking of quantities" in Section 2.6.4.

2.1 Explaining and conceiving

To explain the facts of observation, their occurrence and their recurrence, has been said to be the "distinctive" and "one of the foremost" and even the sole aim of empirical science (Nagel 1961, p. 15; Hempel 1965, p. 245; Popper 1972, p. 191).[1] According to a philosophical tradition that issues from John Stuart Mill's *System of Logic* but can be traced to earlier sources, a scientific explanation takes the form of an inference whose conclusion describes the fact or facts to be explained (the explanandum), while its premises (the explanans, i.e., 'that which explains') consist of the statement of a law of nature and the description of some other facts. This idea of explanation as inference, carefully articulated within logical empiricism (Hempel and Oppenheim 1948; Braithwaite 1953; Hempel 1965), was relentlessly criticized from different standpoints in the sixties, when that philosophical movement, which like Bauhaus architecture and Comintern politics had posed as definitive, turned out to be even more ephemeral than such worldly fashions (Scriven 1958, 1962; Toulmin 1961; Feyerabend 1962; Bromberger 1966; Harré 1970, etc.). I do not intend to repeat here those criticisms or the replies they elicited but rather to concentrate on one feature of deductive explanation which, to my mind, contains the clue to its significance and yet has rarely been in the limelight of philosophical debate.

A deductive explanation, better known as a deductive-nomological or DN explanation (the term 'nomological' being built from νόμος, the Greek word for law), infers the statement of an observed fact F from the joint statement of a law or laws L and of factual conditions C.[2] For the explanation to work, each of the statements L, C, and F must meet certain requirements that need not concern us.[3] F can be inferred from L and C if and only if the conditional $(L \supset (C \supset F))$ is a logical truth. However, the conditional $(C \supset F)$ ought not to be one, or the law L would be superfluous. Therefore, the fact under consideration must be described by F in terms that also occur in L. Typically, the law L will link the terms descriptive of the fact to be explained F with those descriptive of the factual conditions C. We may express this by saying that in DN explanation the explanandum and the law in the explanans must be conceptually homogeneous. This principle is readily illustrated by the following example, long a favorite in philosophy classes. I can infer—and thereby supposedly explain—the fact that

(F^*) This thing here is black

from the general law

(L^*) All ravens are black

and the known condition

(C^*) This thing here is a raven

But the inference from L^* and C^* to the observed fact will not go through if I grasp this thing here as being warm or winged or noisy but do not grasp it as black—as may well be the case if it happens to be a raven that I touch or hear but do not see.[4]

Although the foregoing example meets the stated conditions for DN explanation and clearly illustrates the requirement of conceptual homogeneity, it will probably not be recognized outside philosophical circles as an instance of scientific explanation. Explanation, in the ordinary meaning of the word, should be enlightening; yet the uniform blackness of ravens throws no light at all on the fact that this raven here is black. If we turn to a more likely example—e.g., if we derive the current generated by a particular alternator when it rotates with a given frequency from classical electrodynamical laws and suitable factual conditions—we shall see that it differs from the black raven case in at least the following two respects:

(i) The laws adduced for explanatory inference in real science normally involve concepts alien to prescientific discourse. In order to achieve conceptual homogeneity the facts of observation which are to be explained must somehow be grasped under those same concepts.

(ii) Such strictly scientific concepts are always part of a coherent and explicit—or, at any rate, progressively self-explicating—system of thought that links the explanandum through the laws in the explanans to a variety of other facts, derivable from the same or related laws.

These noteworthy features of standard scientific explanatory inferences are not independent of one another. If natural philosophers and scientists had remained content with the stock of notions of prescientific common sense, instead of developing novel intellectual systems, they would never have been able to bring together such seemingly disparate phenomena as falling apples, orbiting satellites, and receding galaxies, and to have each of them illuminate the others and bestow relative necessity upon them. By bringing their innovative thought to bear on the facts of observation they have succeeded in producing explanatory inferences that truly increase our understanding. One will admit to having understood why some particular thing or event is as it is if one gets to see that it could not be otherwise. Such physical necessities are relative, not absolute, inasmuch as they depend on what the rest of things and events is like. By grasping different facts under concepts bound together in a system, we achieve just this kind of understanding: unless each explanandum within the scope of the system follows from the relevant laws and a suitable description of the prevailing circumstances, we must rethink all other facts within that scope. When a collection of facts is thus incorporated into an intellectual system, each one of them is, so to speak, held in place by the rest. While we grasp it as we do, we cannot conceive it to be otherwise than we think it is, unless we reconceive the other facts in the collection as well. Inferential explanations in genuine science thus differ sharply from the classroom example proposed above. If albino ravens are found in Alaska, we should not feel compelled to rethink our zoology. And the logical necessity with which the conclusion "This thing here is black" (F^*) follows from the premises L^* and C^* does not create even a mirage of physical necessity with regard to the fact F^* itself. In the light of the classroom explanation, this thing here could just as well be green *and* a raven, and all other ravens remain unchanged. On the other hand, should we ever establish with reasonable certainty that a particular planet does not obey the accepted law of gravitation, we would have to revise our thinking about gravitational phenomena throughout the universe.

2.2 Examples from Newton

In order to see better how a systematic rethinking of facts is at work in scientific explanations, we shall now consider a few applications of Newton's Law of Gravity.

Take the motion of the Moon around the Earth. On a first approximation we ignore the presence of the Sun and other heavenly bodies, and we treat the Earth as fixed. The explanation of lunar motion by Newton's Law of Gravity rests then on the assumption that the Moon is freely falling towards the Earth in accordance with that law. Although this is nowadays a trite commonsense idea, it was far from being one in the 17th century. Indeed it must have seemed paradoxical, inasmuch as the falling Moon never reaches the ground. And yet, unless we conceive of the Moon in some such way it would be madness to try to infer a statement of its several positions from a law of gravity.

Newton's conception of the Moon as a falling body derived some plausibility from Galileo's analysis of the motion of projectiles near the surface of the Earth. According to a view current in Galileo's day, a heavy body such as a cannonball will naturally move downwards to the center of the Earth if it is not stopped. However, by force it can be made to move unnaturally upwards or sideways in any direction. But no body will move both naturally and against its nature at the same time. Hence, a cannonball, after being shot, will first be driven by the exploding gunpowder in the direction in which the cannon points, and only when the force of the explosion is spent will it fall—vertically—on its target.[5] This entails that the range of a cannon is greatest when it points horizontally, a prediction not confirmed by experience. Galileo dismissed the assumption that different motions cannot coexist in the same body and chose to think of a cannonball as falling freely from the moment it left the cannon's muzzle. According to Galileo's ideas about free fall, this entails that a flying cannonball suffers at all times the same downward acceleration, regardless of the material of which it is made. Galileo showed by a clever calculation that, on the stated assumptions, a projectile issuing from a horizontal cannon on top of a parapet describes a parabola. If Galileo's principle of inertia is extended to nonhorizontal motion, the result holds also for projectiles shot in any direction.[6]

Galileo's calculation presupposes that the acceleration of gravity is constant in both magnitude and direction. But, of course, if it points to the center of the Earth it can keep steady only within a small region—namely, that within which the Earth's surface may be regarded as approximately flat. Newton assumes, moreover, that its magnitude is the same only at equal distances from the center of the Earth and varies as the inverse square of that distance. It can then be shown that the projectile's trajectory is a conic section, generally an ellipse or a hyperbola. If we think of the Moon as such a projectile and let $\mathbf{v}(t)$ stand for its velocity at a given time t, while $\mathbf{r}(t)$ denotes its position at that time, referred to the Earth's center $\mathbf{0}$, we can readily calculate the acceleration $\mathrm{d}\mathbf{v}(t)/\mathrm{d}t$ if we are given, say, the values ρ and g of the radius of the Earth and the acceleration of gravity at the poles. For then it follows from the said assumption of Newton's that

$$\frac{\mathrm{d}\mathbf{v}(t)}{\mathrm{d}t} = \mathrm{g}\,\frac{\rho^2}{\mathbf{r}^2(t)}\,\frac{\mathbf{r}(t)}{|\mathbf{r}(t)|} \tag{1}$$

The value calculated from this equation agrees passably well with the acceleration needed to account for the observed motion of the Moon.[7]

Newton's Law of Gravity is far bolder and more speculative than the modest and fairly straightforward extension of Galileo's Law of Free Fall that I have sketched here. Yet even within the narrow bounds in which I have deliberately kept our example we can readily see how the familiar data of observation must be rethought before explanatory inference can do its job. Not only must the Moon be conceived on the analogy of a cannonball, but its motion must be described under novel concepts of time, space, velocity, and acceleration, whose systematic interconnections provide the means for comparing and coordinating the lunar data among themselves and with the phenomena of falling bodies. Without these concepts, the falling Moon is no more than a suggestive metaphor; but thanks to them the analogy of projectiles takes a precise and pregnant meaning: ballistic trajectories, that of the Moon included, are solutions, under diverse conditions, of the same set of differential equations. (Whence, by judiciously choosing and effecting still other conditions, we have been able to put all those tiny man-made moons into the sky.)

Had Newton been content with extending earthbound Galilean gravity to the Moon in the manner proposed above he would have provided a good example of cautious generalization from proven facts, but his success would have been short-lived. For the trajectory that can be obtained from eqn. (1), though strikingly accurate as a first approximation, still differs noticeably from the one observed.[8] But Newton thought of mutual gravitation as a universal law of matter. He expressed this law in terms of his original concept of impressed force. By definition, "an impressed force is an action exerted upon a body in order to change its state, either of rest or of uniform motion in a straight line" (Newton, *Principia,* Def. IV). According to Newton's Third Axiom or Law of Motion, to every such action there is always opposed an equal reaction, so that "the mutual actions of two bodies upon each other are always equal, and directed to contrary parts." Newton's own statement of the Second Law of Motion indicates that he meant by impressed force what we now call impulse (with the dimension of mass times velocity); but his use of the concept in actual proofs warrants our understanding of Newtonian force as a cause of acceleration and its familiar representation by a vector proportional to the acceleration caused by it (the factor of proportionality being equal to the mass or quantity of matter of the accelerated body).[9] Building on these ideas, Newton attributed the accelerated fall of bodies towards the center of the Earth and the continual deviation of planets from rectilinear motion to an attractive force exerted by every material particle on all the others. By a liberal application of his professedly inductivist methodology he concluded that this force is directly proportional to the mass of both the attracting and the attracted particle, and inversely proportional to their distance squared.[10]

This Law of Universal Gravitation furnished Newton and his successors with an extremely supple and efficient instrument for calculating planetary motions. Yet at first blush it might seem to raise an unsurpassable difficulty. It is a truism often forgotten in empiricist discussions of empirical science that the data collected by observation cannot be explained by inference from general laws unless they have been rendered comparable. This requirement is met in Newtonian physics by referring all data on matter and motion to a common space and time. Now, Newton's "absolute, true, and mathematical" time and space cannot be observed but must be constructed from the relative times embodied in mechanical clocks and the relative spaces sustained by material frames of reference. But if every speck of matter is continually being pulled in every direction by all the rest, any reference frame or clock one may chance to choose is likely to be accelerated in a wholly unpredictable way. How can one expect to gather in a single coherent system of comparable kinematic data the results of observations referred to frames and clocks whose true state of motion is unknown? This difficulty, however, is satisfactorily resolved by invoking two principles inherent in Newton's basic kinematic and dynamic assumptions, that is, in his concepts of space and time and in his Laws of Motion. By the justly celebrated Newtonian Principle of Relativity (Corollary V to the Laws of Motion),

> The motions of bodies included in a given space are the same among themselves, whether that space is at rest, or moves uniformly forwards in a straight line without any circular motion.

This means that an inertially moving frame and a good mechanical clock affixed to it are adequate substitutes for absolute space and time (provided that one assumes, with Newton, that "every moment of time is diffused indivisibly throughout all spaces"—Hall and Hall 1978, p. 104). Yet in a world held together by universal gravitation one is hard put to find a body in true inertial motion. By the no less significant Newtonian Principle of Equivalence (Corollary VI to the Laws of Motion),

> If bodies, moved in any manner among themselves, are urged in the direction of parallel lines by equal accelerative forces, they will all continue to move among themselves, after the same manner as if they had not been urged by those forces.

Thanks to this principle, the Newtonian astronomer need not be worried by the farfetched but not impossible thought that the entire firmament of the fixed stars may be gravitating towards a remote and invisible but enormous concentration of matter. Indeed he may comfortably ignore the circumambulation of the Sun in the Galaxy and of the Galaxy in the Local Group—not discovered until the 19th and the 20th century, respectively—and carry out his investigation of planetary motions as if the center of gravity of the solar system were at rest. Moreover, he may neglect, on a first approximation, as we did above, the conspicuous acceleration of the Earth towards the Sun and treat the Earth-Moon system as if it were isolated. For during the short intervals required for ascertaining the velocity and acceleration of the Moon in the relative space of the Earth, the force exerted on the system from the Sun does not vary appreciably in magnitude or direction. (See Stein 1977, pp. 19ff.)

We thus see how radically committed to the Newtonian mode of thought is even the humblest explanation of an observed fact by Newton's Law of Gravity. The reason for this is plain enough. To be covered by the Law the fact has to be embedded in the structure of Newtonian kinematics: the times of observation must be instants—or very short intervals—of universal Newtonian time, the observed positions must be located in an admissible Newtonian relative space, measured distances must satisfy the applicable theorems of Euclidian geometry, velocities must behave as smooth vector-valued functions of a real variable. Moreover, Newtonian kinematics is inextricably intertwined with dynamics (see, e.g., Torretti 1983, Chapter 1). In contrast with such hackneyed "laws" as "All ravens are black," or "Water cleans," which demand little by way of intellectual commitment, the laws of mathematical physics, exemplified by Newton's Law of Gravity, are always deeply involved with some exacting theoretical system, apart from which they have no definite meaning. While empirical generalizations of the former kind can hardly be said to explain anything, a theory-laden law, such as Newton's, has great explanatory power thanks to the links induced—or should we rather say disclosed?—in a rich variety of facts by embedding them in the law's underlying theoretical structure.

The intellectual efficacy of such links can be further clarified by considering a case in which Newton's Law of Gravity does not provide a satisfactory explanation of facts presumably within its scope. Mercury's perihelion advances each year by somewhat less than 1 minute of arc. 90% of this advance—some 5,000″ per century—is due to the precession of the axis of the Earth, to which astronomical coordinates are referred. An additional 9% can be accounted for by the action of the other planets in agreement with Newton's Law (280″ per century can be ascribed to the action of Venus, 150″ to that of Jupiter, 100″ to the rest). But there remains a balance of approximately 43″ per century which, under the known circumstances, cannot be explained by Newton's Law. If the planet Vulcanus, invented for just this purpose, had been discovered or if the oblateness of the Sun (i.e., the ratio of its equatorial diameter to the distance between its poles) were significantly larger than it appears to be or if the dust surrounding the Sun were dense enough, the small secular anomaly of 43″ would be readily covered by Newton's Law. On the other hand, the anomaly agrees uncannily well with the prediction of Einstein's theory of gravitation, designed, as against Newton's, to fit the new Einsteinian conceptions of space and time. Since Einstein (1915) made this agreement known, the anomalous advance of Mercury's perihelion has come to be generally regarded as one of the classic instances in which the Newtonian explanation of planetary motion fails.[11] Now, the interesting thing to note is that, when one takes this stance, one cannot simply view Mercury—as one would a green raven—as a doubtless repeatable but on the whole unlikely exception to the otherwise well confirmed ordinary course of nature, but one must conclude that the other planets, whose observed behavior has hitherto agreed well with Newton's Law, follow it only approximately, within a margin compatible with current observational imprecision, due to their particular circumstances.[12] This inference from the failure of Newton's Law in a single case to its universal invalidity is forced on us by the very concepts of mass, space, velocity, etc., by which planets and their motions must be grasped in order to subject them in scientific discourse to that law. It is applicable to them only if they are bodies of gravitationally homogeneous mass moving in homogeneous Euclidian space. If they are such, either Newton's Law is true of them all or the recorded compliance of most planets with it is a mere coincidence. Hence, if one gives up all hope of accounting for a proven anomaly by Newton's Law and some hitherto undetected factual condition (as Leverrier and Adams explained the anomalous motion of Uranus by the Newtonian attraction of the then unknown planet Neptune), there are only two viable ways of dealing with it: either one tries a different law conceived in the same terms as Newton's—e.g., one in which the factor r^2 has been replaced by a less simple function of the distance between the interacting particles—or one builds upon a different conceptual foundation and comes up with a different understanding of gravitational phenomena.[13] But under no circumstances can one maintain, in the face of a single avowedly insoluble anomaly, that "Eight planets out of nine attract each other (and the Sun) directly as their masses and inversely as their distances squared"—as one may still endorse "Water cleans" even while trying in vain to wash out with water a stain in one's clothes. Systematic thought breeds necessity in such a way that, when the latter is found wanting, the system itself loses its hold on things.

2.3 Questions raised by conceptual innovation

Examples similar to those of Section 2.2 can be found in all fields of mathematical physics. They tend to show that—at any rate in this branch of science—explanatory inference is only a step or a facet of a process of thought whose decisive stage consists in producing concepts appropriate for grasping a variety of facts and linking them together in an intellectual system. Systematic linkage of many ostensibly diverse facts serves to corroborate the appropriateness of the concepts by which each of them is grasped, and is indeed the mainstay of inferential explanation. In Chapter 3 I shall have more to say about such intellectual systems. But let us first consider a question that one is bound to face sooner or later if it is true that facts of observation are grasped and regrasped under changing concepts.

Somewhat schematically that question can be introduced as follows: Science responds to puzzling facts, which it seeks to explain by reconceiving them. The facts are embedded in a conceptual system within which they follow from general laws, given their particular circumstances. In the course of this process, science discards the original description of the facts, which set the inquiry in motion, and proposes a new description, under which they are no longer puzzling. To what extent and by virtue of what device does the new description, required for the proposed explanation to work, refer to the same facts as the old description, which caused an explanation to be sought for them?

More pointedly, we may ask:

Q1. Can the facts remain the same as the framework of description varies?

Q2. Does the apparent need for a steady reference set some permanent limits to conceptual innovation?

Q3. Is such a steady reference really necessary, or may the researcher, without detriment to the rationality of his enterprise, sometimes forget along the way the facts he originally had in mind?

The breakdown of reference due to conceptual innovation is a recurrent theme in the historicist school of philosophy of science initiated in the fifties by Norwood Russell Hanson and Paul K. Feyerabend. Hanson bids us imagine the 16th century astronomer Tycho Brahe, a firm believer in the fixity of the Earth, and his Copernican assistant, Johannes Kepler, as they watch the dawn from the top of a hill. Hanson asks, *"Do Kepler and Tycho see the same thing in the east at dawn?"* (Hanson 1958, p. 5). He argues that, though "Tycho and Kepler are both aware of a brilliant yellow-white disc in a blue expanse over a green one," they cannot properly be said to witness the same fact. For "Tycho sees the sun beginning its journey from horizon to horizon. He sees that from some celestial vantage point the sun (carrying with it the moon and planets) could be watched circling our fixed earth. [. . .] But Kepler will see the horizon dipping, or turning away, from our fixed local star" (Hanson 1958, pp. 7, 23).

In a series of papers, Feyerabend (1958, 1960, 1962, 1965) repeatedly emphasized that terms inevitably change their meaning as they pass from the context of one scientific system into that of another. It is therefore very difficult to compare how well several such systems "fit the facts." A valid comparison is downright impossible when the systems under consideration concern the basic elements and properties of the universe. Each system will then "possess its own experience, and there will be no overlap between these experiences. [. . .] A crucial experiment is now impossible [. . .] because there is no universally accepted *statement* capable of expressing whatever emerges from observation" (Feyerabend 1965, p. 214).

Similar ideas were voiced, somewhat fuzzily, but with great rhetorical efficacy, by Thomas S. Kuhn (1962). To him a scientific revolution involves "a displacement of the conceptual network through which scientists view the world" (1962, p. 101). As a consequence of such a displacement, the new scientific tradition that issues from a scientific revolution "is not only incompatible but often actually *incommensurable* with what has gone before" (1962, p. 102; my italics). Therefore, one "may want to say that after a revolution scientists are responding to a different world" (1962, p. 110).

Kuhn's strong claims concerning the incommensurability of alternative modes of scientific thought and the substitution of one world for another in the course of a scientific revolution are clearly unjustified in those cases where enough remains of the prerevolutionary conceptual setup to allow a shared description of crucial facts. Thus, for example, deep though it was, Darwin's revolution in biology did not affect the distinction between living organisms and inanimate bodies nor such descriptions of the structure and behavior of the former as may be adduced for resolving the dispute between Darwin and his adversaries. Hence, someone who does not share Darwin's vision may dismiss the evidence gathered in *On the Origin of Species* as inconclusive, but not as unintelligible, as he might do if it were expressed in esoteric terms peculiar to the doctrine he rejects, and not in plain English.

It would seem, however, that reference to the selfsame facts will not survive conceptual renewal when this involves the very notions in terms of which the phenomena of motion and the states of physical systems are described. Yet even in this case, we can think of three conditions any one of which is sufficient to ensure the comparability of scientific claims in the face of such radical conceptual innovation. The continuity of scientific discourse can be preserved if:

C1. Some concepts are immune to change, and they provide a stable reference to decisive facts.

C2. The new concepts are arrived at through internal criticism of the old, by virtue of which the facts purportedly referred to by the earlier mode of thought are effectively dissolved.

C3. Reference to facts does not depend on the concepts by which they are grasped.

These three conditions are closely linked to the three questions regarding conceptual innovation I raised earlier in this section. Thus, an affirmative answer to question Q2 entails condition C1. Condition C2 would warrant an affirmative answer to the second part of question Q3. Finally, even if Q2 were to receive a negative answer, the fulfillment of C3 would justify an affirmative reply to Q1: if reference does not depend on concepts it may very well remain steady even when concepts change. In the next two sections I shall examine C1 and C2 in the tacit understanding that C3 does not hold. Then, in Section 2.6 I shall argue that, notwithstanding recent allegations to the contrary, C3 must be denied.

2.4 Are there limits to conceptual innovation in science?

2.4.1 Self-classifying sense impressions

Conceptual innovation will be confined within definite limits if all our experience of the world is compounded by association from simple, repeatable, self-classifying sense impressions. Scientific concepts would then represent different kinds of combinations or combinations of combinations, etc., of such simple impressions, and the vocabulary of science would fall into two parts:

(i) A basic vocabulary V_O, each term of which would designate one of the known classes of simple sense impressions or the simple relation of association between them

(ii) A derived vocabulary V_T, whose terms would signify the several ways of combining those sense impressions or their combinations or combinations of combinations, etc., to any order, which are felt to merit a label

The extension of each term of V_O would then be fixed once and for all by the natural self-classification of sense impressions. The list of such terms could only grow or decrease together with our capacity for receiving different sorts of impressions, and would therefore be stable, except during periods of major change in the genetic makeup of man. Innovation would be confined to terms in V_T and the concepts they express, which every scientist would indeed be free to fashion and refashion at his pleasure. Such terms, however, would be completely meaningless unless connected by definitions or other meaning-bestowing devices with terms of the basic vocabulary V_O. Any descriptions of particular facts involving terms in V_T would be replaceable, without prejudice to truth, by equivalent descriptions of the same fact that use only terms in V_O. The latter descriptions would anyway be shared by alternative systems of scientific explanation, in spite of any differences in their V_T vocabulary. Thus, it would not be impossible to compare them and to ascertain which one among them provides more appropriate premises for inferring the description of any given fact observed.

The scheme just sketched is, of course, chimerical. In real life, private sense impressions, far from providing the ultimate foundation of all experience, turn up—e.g., at the ophthalmologist's or while tasting wines in a winery—only in settings firmly anchored to public physical objects. Moreover, they are never simple, and they display in and of themselves no indication as to how they ought to be classified. Thus, for example, nothing in the sheer visual appearance of a rainbow could constrain us to see just the six colors we normally distinguish in it, instead of three (with Aristotle) or seven (with most nursery school teachers). Indeed, as the same example suggests, any classification of sense appearances is open to refinements and displacements due not to genetic mutations but to our changing interests and attention.

2.4.2 Kant's forms and categories

A very different treatment of the question of concept stability can be extracted from Kant's *Critique of Pure Reason*. Kant was firmly committed to a conception of man as a "finite" subject of knowledge, who can only learn about an object by the way the object "affects" him. But Kant saw clearly that a knowledge of objects could not result from the mere association of subjective affections. The core of his book is an inquiry concerning the "conditions of possibility" of our human experience of the physical world. He divides them into two classes. There are, in the first place, the "forms" of sense awareness, which make it possible that the manifold of sense appearances be ordered in certain relations.[14] These forms he identifies with time and space, which he regards as inherent conditions of our "receptivity" to sense impressions. In the second place, there are the "functions" by which our understanding combines and unifies the given manifold of sense in such a way that it is construed as a presentation of objects.[15] The "categories," or fundamental concepts of ontology, exactly correspond to the said "functions" of the understanding "insofar as the manifold of a given intuition is determined with respect to them" (Kant 1787, p. 144). Kant derives an allegedly complete list of the "categories" from the classification of "judgments" found in contemporary textbooks of logic, suitably enriched with two unfamiliar items, viz., "singular" and "infinite" judgments, to meet the desiderata of ontology. He claims that this classification reflects the several functions of the understanding. He argues, more convincingly, that temporal self-awareness presupposes awareness of enduring objects in space. He takes it for granted that the ordering of sense appearances in relations of space and time necessarily complies with the principles of Euclidian geometry and Newtonian chronology. He contends that every distinct content of sense awareness (every shade of color, tone of sound, etc.) must be grasped as belonging to some continuous scale of intensities that goes from the presence of that qualitatively peculiar content right down to its total suppression, passing through every conceivable intermediate degree. In a long chapter on "The Analogies of Experience" he offers proof that the objective time order of phenomena—as opposed to the merely subjective succession of appearances[16]—can only be established by grasping them under the categories of "subsistence and inherence (substance and attribute)," "causality and dependence (cause and effect)," and "community (reciprocity between agent and patient)," subject to the principles of conservation of the quantity of matter, causal determinism, and thoroughgoing instant interaction. The "forms" of time and space, the categories of the understanding, and the principles that govern the application of the latter to the manifold displayed in the former are, according to Kant, permanent features of human reason, such that "we cannot form the least conception" of a cognitive faculty which worked differently. We must, therefore, deem them necessary, although we cannot give any grounds "why we have just these and no other functions of judgment, or why space and time are the only forms of our possible intuition" (Kant 1787, p. 146).

Kant says that the "unity of consciousness" built by exercising the "functions" of the understanding on the manifold of sense "is that which alone constitutes the reference of representations to an object" (Kant 1787, p. 137). If this is granted, the Kantian system of categories and "forms" can certainly fix the reference of our factual descriptions in the face of conceptual innovation. For such innovation can then concern only empirical concepts that do no more than specify the categories, and it will be constrained by the principles of the understanding that preside over the articulation of sense appearances into an experience of physical objects. Reference to *the same thing* can be secured, in this view, regardless of any changes in its attributes or in the way we describe them and classify them, by following a constant "quantity of matter" in time as it moves in space. Thus, going back to the examples of Section 2.2, we can see at once that, if we refer to the several components of the solar system by giving the position at each instant of their respective masses, we fix thereby all the facts that any dynamic theory of planetary motion, consonant with Kant's philosophy, should seek to account for.

One could still object that the method of identifying bodies by timing (in Newtonian time) the positions (in Euclidian space) of their real-valued masses was inaugurated by Newton and his contemporaries, and therefore could not secure the stability of reference at the transition from pre-Newtonian to Newtonian physics. But a Kantian may well counter this complaint by recalling that "the highway of science" ("der Heeresweg der Wissenschaft"—Kant 1787, p. xii) has been entered upon by each branch of inquiry at a certain point in history; and that only from then on—i.e., in the case of physics, only from the 17th century (Kant 1787, p. xii)—can it boast an objective representation of phenomena, which both sets the task and provides a test for alternative scientific explanations of them.

A much more damaging criticism of the Kantian position results from considering the actual scope of conceptual novelty in the two major systems of physical thinking that have replaced Newton's in the 20th century, namely, Relativity and Quantum Theory. The former took issue with Newtonian physics on the selfsame notions of time, space, and mass that were for Kant the key to objective reference; the latter gave up the principle of causal determinism, without which, according to Kant, there could be no experience of objective succession.[17] As to the principle of thoroughgoing instant interaction, which he had considered indispensable for establishing the objective simultaneity in space, Relativity dismissed it from the outset. Thus, in the first major revolutions in mechanics after the publication of Kant's book, his system of "forms," categories, and principles could not stem the tide of conceptual innovation but was swept away by it.

We cannot go further into this matter here (though I shall have something to say on Einstein's criticism of Newtonian time in Section 2.5 and on his several concepts of mass in Section 2.6). I take it that the bare mention of those theories is sufficient to remind us that the Kantian restrictions on conceptual innovation have proved untenable. Some philosophers believe that Kant's "metaphysics of experience" was too strong and trespassed on matters pertaining to empirical science but that if only it is conveniently weakened, it can and must be upheld (Rosenberg 1980; Stevenson 1982). I shall touch again on this issue in Section 2.7. But here I wish to consider another way of stabilizing factual reference by restricting the admissible range of conceptual novelty, which was developed in the second third of the 20th century, mainly by Rudolf Carnap.

2.4.3 Carnap's observable predicates

In *Der logische Aufbau der Welt* (1928), Carnap proposed a method of "reducing" all objects—in the widest sense, including things, states and events, properties and relations—with the aid of modern logic, to a "basis" of homogeneous "ground elements" and fundamental relations between them.[18] Carnap admitted the possibility of adopting a physical basis with one of the following alternative sets of ground elements: (α) electrons and protons, (β) spacetime points, (γ) point-events on the worldlines of matter.[19] He chose, however, a solipsistic psychical basis, whose ground elements are the *Elementarerlebnisse,* or instantaneous cross sections of the total stream of a person's mental life.[20] Following Nelson Goodman (1966, p. 154), I call such ground elements *erlebs.* As fundamental relation Carnap chose *Ähnlichkeitserinnerung* (literally, 'recollection of resemblance'), i.e., the relation between two erlebs x and y that are known to be similar by comparing y to a memory of x. Carnap's preference for this basis was motivated by his conviction that knowledge of one's own stream of erlebs was presupposed by one's knowledge of anything else. Reduction to a solipsistic basis was therefore required in order to be faithful to the epistemic hierarchy of objects.[21] That every object of science can be reduced to erlebs he proved as follows:

> If any physical object were not reducible to sensory qualities and hence to psychical objects, that would mean that there are no perceptible criteria for it. Statements concerning it would then dangle in a void. At any rate, it would have no place in science.
> (Carnap 1961, p. 78)

The reductive program of Carnap's *Aufbau* has time and again enticed talented writers to devote their energy and ingenuity to its advancement (Goodman 1951; Moulines 1973). However, Carnap himself moved further and further away from it, dropping its main assumptions one by one. Yielding to Otto Neurath's criticism he substituted a "physicalistic" basis for the solipsistic erlebs (Carnap 1932). Later, he liberalized the requirements of reduction, to make allowance for scientific talk of dispositions that might never be actualized—such as the solubility in water of a substance that will never be removed from a waterless planet (Carnap 1936/37). Finally, he gave up the very idea of reduction and replaced it with a program of "partial interpretation" (Carnap 1956). Yet he did not relinquish his dream of science anchored forever to the rock bottom of the epistemic hierarchy. It is indeed ironic that he should have once sought to build his cathedral of knowledge on the drifting sands of *Erlebnis*. But when he opted for physicalism, the basis of spacetime points he originally countenanced was soon discarded, for it clearly did not enjoy epistemic primacy. Speaking the "formal idiom"—talk of words—which he now preferred to the more familiar but potentially misleading "material idiom"—talk of objects—he had used before,[22] he demanded that scientific discourse be reduced to *thing-language*— i.e., "that language which we use in every-day life in speaking about the perceptible things surrounding us" (Carnap 1936, p. 466)—and, more specifically, to the "observable predicates of the thing-language" (1936, p. 467). The key notion of an *observable predicate* Carnap explicated as follows:

> A predicate 'P' of a language L is called *observable* for an organism (e.g. a person) N, if, for suitable arguments, e.g. 'b', N is able under suitable circumstances to come to a decision with the help of a few observations about a full sentence, say 'P(b)', i.e. to a confirmation of either 'P(b)' or '~P(b)' of such a high degree that he will either accept or reject 'P(b)'.
>
> (Carnap 1936, pp. 454ff.)

He noted that "there is no sharp line between observable and non-observable predicates because a person will be more or less able to decide a certain sentence quickly." However, "for the sake of simplicity," he chose to draw a sharp distinction—"in a field of continuous degrees of observability"—between observable and non-observable predicates (Carnap 1936, p. 455).[23]

Twenty years later, Carnap (1956) still held to this sharp distinction but acknowledged that some nonobservable terms of science may be irreducible. He took for granted that scientific discourse is expressed in a formal language—or in a readily formalizable segment of a natural language—which he called "the language of science," L. This language falls neatly into two parts, the "observational language" L_O, and the "theoretical language" L_T. L_O is an interpreted first-order language,[24] whose variables range over "concrete, observable entities (e.g. observable events, things, or thing-moments)" and whose predicates—the "observational vocabulary" V_O—designate "observable properties of events or things (e.g. 'blue', 'hot', 'large', etc.) or observable relations between them (e.g. 'x is warmer than y', 'x is contiguous to y', etc.)" (1956, p. 41). L_T is a predicate calculus that *may* include logical and causal modal operators, besides the usual quantifiers and truth-functional connectives. The domain D over which the variables of L_T may range in a given interpretation is subject only to the following conditions: (i) D includes a distinguished countable subdomain, and (ii) D is closed under the operations of n-tuple formation (for every positive integer n) and class formation (1956, p. 43). Let T denote a finite set of sentences of L_T. The set of the logical consequences of T is a "theory," also designated by T, for which T, the finite set, provides the postulates.[25] The predicates of L_T which occur in T (in either meaning) form the "theoretical vocabulary" V_T of the theory. Carnap believes that any meaningful scientific use of "the language of science" L involves a choice of such a theory T.[26] According to him, any such use also requires the stipulation of a set C of "correspondence rules," which license the drawing of conclusions in L_O from premises in T (usually conjoined with premises in L_O). T owes its cognitive meaning to the consequences in L_O which thus accrue to it through C. Its empirical support depends on their truth. The correspondence rules C take the form of additional postulates or of rules of inference. In either case, their formulation in L must include predicates from both V_T and V_O. But not every predicate in V_T must occur in C. It is enough that the correspondence rules link *some* of the theoretical predicates to observational predicates. The interpretation of theoretical discourse by its observable consequences need only be partial.

Carnap's theory of scientific "theories" was not designed to cope with the problem of scientific change—to which, indeed, its author was notoriously insensitive. But it may be readily adapted for that purpose. Restrict the innovative action of thought to the theoretical vocabulary of science and the postulates and correspondence rules in which it is embedded. Then, the observational vocabulary, changing, if at all, at a much more leisurely pace, undisturbed by the vicissitudes—and the achievements—of theory, provides the stability of reference required for a comparison between the observable consequences of any two rival theories. Since Carnap and his followers issued no disclaimers when they were criticized by Feyerabend and Hanson for just this type of restriction, one may conclude that they approved it. And yet one cannot but feel astonished at the sheer extravagance of supposing that the language in which scientific observations are reported should be out of bounds for scientific thought.

I shall not review the arguments against a distinction between observational and theoretical terms in science. If the reader is not acquainted with them, I suggest reading Hilary Putnam's "What Theories Are Not" (1962) and Mary Hesse's "Is There an Independent Observation Language?" (1970). After two decades during which the distinction was unfashionable (but see, however, Shimony 1977, which is the printed version of a lecture dating from 1969), trendy writers are now saying that there was something to it. Ian

Hacking (1983, p. 175) pokes fun at the notion that all terms are "theory-laden"—and quite rightly, I dare say, for surely we do not wish to claim that when Monsieur Jourdain shouted, *"Nicole, apportez-moi mes pantoufles, et me donnez mon bonnet de nuit,"* his words were loaded with anything we would call a theory. In a more conservative vein, W. H. Newton-Smith, while dismissing "the alleged O/T dichotomy," sponsors

> a rough and ready differentiation between the more observational and the more theoretical [. . .] determined by the following principles:
>
> 1. The more observational a term is, the easier it is to decide with confidence whether or not it applies.
> 2. The more observational a term is, the less will be the reliance on instruments in determining its application.
> 3. The more observational a term is, the easier it is to grasp its meaning without having to grasp a scientific theory.
>
> (Newton-Smith 1981, pp. 26–27)

As a rough and ready characterization, the above is neat enough. But are the three stated criteria mutually consistent? To avoid disputes about the meaning of 'meaning' I leave the third one aside and concentrate on the first two.[27] Many familiar words score well on both counts. But confidence in the use of a term is not necessarily stronger because one does not rely on instruments in determining its application. Having just weighed 263 ± 0.2 grams of corn flour in a precision balance I am certainly more confident that it weighs 263 grams (to the nearest gram) than that it is corn flour. Roughly speaking, the terms that we use most confidently have to do with the practice of life—including, of course, the life of science—and their applicability can be tested in action. I rest assured that it is an axe I hold in my hands if I can chop a log with it. But our confidence need not decrease because the practical test involves the use of instruments. I am certain that power is back after a blackout as soon as I hear music on my tuner; I do not have to test my belief by plugging my fingers into an outlet. There is, however, one common circumstance that could perhaps suggest that our confidence in the application of terms is greater, the less we rely on instruments for deciding it. Many ordinary terms are employed very confidently just because the standards for applying them are quite loose. Such terms are not usually the sort whose application is controlled by means of instruments. But neither are they a source of "cognitive meaning" for the theoretical language of science.[28]

A major obstacle to the ordering of scientific terms from the more observational to the more theoretical—and, a fortiori, to their partition into two classes—is that the same term often functions, in different contexts, at either end of the proposed scales. For example, the term 'free fall', as instantiated by a falling stone, is highly observational by all three of Newton-Smith's criteria. But the same term is applied to the motion of the Moon on the strength of the Newtonian or the Einsteinian theory of gravity and the many instrument-assisted observations of planets, pendula, the Moon itself, etc., which corroborate those theories. In this use, therefore, the term is very theoretical by Newton-Smith's criteria 2 and 3, even though, after all the successful experimenting with artificial satellites in the last twenty years, it surely is very observational by criterion 1. The astronomical and the terrestrial uses of the term are not just homonymous, nor are the former only a metaphorical extension of the latter. (As a metaphor, it would be a rather poor one.) When Newton conceived the Moon as a freely falling body, he at the same time implied that cannonballs were but slow moons. His bold thought changed both the denotation and the connotation of 'free fall'. We are now quite certain that he was substantially right, that the term applies in its new sense both to its new and to its old extension, because we know from practice that a marginal increment in energy can convert an earthbound missile into a heavenly body.

As the preceding example shows, there is a legitimate distinction between the more familiar and the more technical uses of language, between everyday words and terms of art (see Wittgenstein, BB, p. 81; PU, §18). But words move back and forth from one category to the other, and—what is more important for our philosophical discussion—scientific usage, once established, claims—and gradually achieves—a controlling role over ordinary language. 'Distance', 'force', 'heat', 'light' are everyday words that modern physics has reclaimed and made precise; and we would not dream of using them, except metaphorically, in a way incompatible with their technical meaning. 'Gas' and 'electricity', first introduced as terms of art, have become kitchen words, but physics is still the acknowledged keeper of their primary meaning. Who is to be the master is less clear in the case of a term like 'energy', originally invented for strictly technical use within a system of thought that is no longer accepted. Modern physics employs it in her own way—different from Aristotle's—but she has not been able to inhibit or regulate its use in journalism and pseudoscience. Physical terms of art naturally belong to theories and are generally applied with the aid of instruments. They are therefore "less observational" by Newton-Smith's criteria 2 and 3. Whether they are so also by criterion 1 will depend mainly on the trustworthiness of the relevant theory. Misologists, to whom anything that smacks of intellect is suspicious, believe that familiar words, like 'soap' and 'clean', can be employed more confidently than terms of art, like 'entropy' or 'inductance'.

Yet if engineers were often wrong in their applications of the latter, the costs would be unbearable.

Clearly, it is very difficult—perhaps impossible—to isolate a family of meaning-invariant English words that regularly satisfy Carnap's criterion for observable predicates or Newton-Smith's criteria for "more observational" terms. No wonder that Carnap and his followers seldom give any examples of the "correspondence rules" by which scientific discourse is supposedly anchored to the observational vocabulary, and when they produce one it leaves much to be desired. No such example is to be found in Carnap's paper of 1956; but in his "Foundations of Logic and Mathematics" (1939) he had proposed the following:

> Let us imagine a calculus of physics constructed [. . .] on the basis of primitive specific signs like 'electromagnetic field', 'gravitational field', 'electron', 'proton', etc. The system of definitions will then lead to elementary terms, e.g. to 'Fe', defined as a class of regions in which the configuration of particles fulfils certain conditions, and 'Na-yellow' as a class of space-time regions in which the temporal distribution of the electromagnetic field fulfils certain conditions. Then semantical rules are laid down stating that 'Fe' designates iron and 'Na-yellow' designates a specified yellow color. [. . .] In this way the connection between the calculus and the realm of nature, to which it is to be applied, is made for terms of the calculus which are far remote from the primitive terms.
>
> (Neurath, Carnap, and Morris 1971, vol. I, pp. 207–8)

Now, if this example is to be taken seriously, one must point out at once that never in the history of chemical nomenclature has 'Fe' designated just any old piece of metal which an ironmonger might accept as iron. Far from being "anchored to the solid ground of observable facts" through the stable reference of the ordinary, "observable" predicate 'iron' as in the stipulation put forward by Carnap, the chemical and physical theories to which the term 'Fe' belongs have set new standards for the application of 'iron', and for the classification, evaluation, and improved production and utilization of what goes by that name. Through its association with 'Fe', 'iron' ceases to be a common everyday predicate, decidable "after a few observations," and becomes a scientific term, whose accurate application relies on appropriate laboratory procedures. Needless to say, if 'iron' had not been thus linked to 'Fe' it would not apply to the gaseous element of atomic number 26 which can be spectroscopically detected in stars. As to the other part of Carnap's example, involving the term 'Na-yellow', I must confess that I do not know what Carnap means here by "a specified yellow color." In ordinary English one may say that a can contains paint of a certain color, e.g., canary yellow, although the paint has never seen the light. But the "temporal distribution of the electromagnetic field" inside a closed can of paint cannot meet the conditions for 'Na-yellow'. One is therefore tempted to believe that, notwithstanding his avowed conversion to physicalism, Carnap understands here by "a specified yellow color" the chromatic quality sensed by a healthy human observer who sees a surface of that color under so-called normal illumination.[29] Now, a scientist who wishes to learn how to recognize Na-yellow radiation at a glance must develop the habit of associating the term 'Na-yellow' with the—presumably stable—color he sees when his eyes receive radiation of that frequency (as measured with the appropriate instruments). One may certainly say that a habit of this sort establishes a semantic rule. The scientist's subjective chromatic experience and any name he may have for it in a private language acquire thereby an objective significance. But the transfer of cognitive meaning effected by such a semantic rule follows a direction exactly opposite to the one indicated by Carnap, viz., *from* the frequency measured in the laboratory *to* the class of erlebs by which the scientist will henceforth diagnose it. That 'Na-yellow' does not simply designate a specific yellow color, in the sense explained, can be inferred from the fact that "the class of space-time regions" that satisfy the requirements for Na-yellow includes some in which the radiation energy is too weak to be recorded by a human eye and some in which it is so strong that it will burn away any organism.

But maybe Carnap's example was only a didactic ploy, not intended for critical scrutiny.[30] To judge the efficacy of correspondence rules one ought then to look for a presentation of a genuine physical theory that actually introduces some of its terms by means of such rules. Carnap's own helpless attempt at formalizing the foundations of Relativity does not reach the point of furnishing an interpretation of the axioms (Carnap 1958, pp. 197ff.). But Hans Reichenbach's *Axiomatik der relativistischen Raum-Zeit-Lehre* (1924) does include what appear to be the equivalent of Carnap's correspondence rules, under the name of coordinative definitions (*Zuordnungsdefinitionen*). The following are typical examples:

Definition 9. Light rays are **straight lines**.
Definition 18. A **natural clock** is a closed periodic system.
Definition 19. A **rigid rod** is a solid rod that is isolated from all external forces.[31]

I have set the definienda in boldface. Each is a term of art of the theory. The light rays mentioned in Definition 9 should of course traverse a perfect vacuum. Thus, this concept is plainly not observational by logical empiricist criteria. Neither is the concept of physical isolation or closure that occurs in the definiens of Definitions 18 and 19.[32]

2.5 Conceptual criticism as a catalyzer of scientific change

Thomas S. Kuhn has noted that "in periods of acknowledged crisis [. . .] scientists have turned to philosophical analysis as a device for unlocking the riddles in their field" (Kuhn 1962, p. 88). In a paper on thought experiments first published in 1964 he explains how a *Gedankenexperiment* proposed by Galileo "helped to teach [. . .] conceptual reform":

> The concepts that Aristotle applied to the study of motion were, in some part, self-contradictory, and the contradiction was not entirely eliminated during the Middle Ages. Galileo's thought experiment brought the difficulty to the fore by confronting readers with the paradox implicit in their mode of thought. As a result, it helped them to modify their conceptual apparatus.
>
> (Kuhn 1977, p. 251)

Kuhn, however, does not sufficiently stress that when a new mode of thought issues from conceptual reform, the problems raised by its purported incommensurability with what went on before are automatically—and trivially—solved. For there can be no question of choosing between two modes of thought if the very existence of the one issues from a recognition of the conceptual failings of the other. If the old is disqualified by the same exercise in criticism that ultimately leads to the new, a comparison between them is not even called for.

The First Day of Galileo's *Dialogo sopra i due massimi sistemi del mondo* contains several fine examples of conceptual criticism, aimed at dislodging the Aristotelian cosmology. The thought experiment to which Kuhn refers is one of them, and here I shall touch on another. Aristotle's system of the world rests on his doctrine about the natural local motion of the elements. Being simple, elements must move simply, unless they are compelled by an external agent to move otherwise. Aristotle recognized two kinds of simple local motion, corresponding to the two classes of lines from which all trajectories are compounded, viz., the straight and the circular. Since the four known elements, earth, water, air, and fire, move naturally in straight lines to and from a particular point, Aristotle concludes that there must exist a fifth element which naturally moves in circles about that same point (*De Caelo*, I, ii–iii; see in particular, 268^b11, 269^aff., 270^b27ff.). This element is the material out of which the heavens are made, and the said point is therefore the center of the world. This is Aristotle's reason for separating celestial from terrestrial physics, and as Galileo's spokesman Salviati points out, it is indeed "the cornerstone, basis, and foundation of the entire structure of the Aristotelian universe" (Galileo, EN, VII, 42). But even granting the premises, Aristotle's conclusion does not follow, for, as Galileo's Sagredo is quick to note,

> if straight motion is simple with the simplicity of the straight line, and if simple motion is natural, then it remains so when made in any direction whatever; to wit, upward, downward, backward, forward, to the right, to the left; and if any other way can be imagined, provided only that it is straight, it will be suitable for some simple natural body.
>
> (Galileo, EN, VII, 40)

Similarly, any circular motion is simple, no matter what the center about which it turns. "In the physical universe there can be a thousand circular motions, and consequently a thousand centers," defining "a thousand motions upward and downward" (Galileo, EN, VII, 40). Salviati goes even further:

> Straight motion being by nature infinite (because a straight line is infinite and indeterminate), it is impossible that anything should have by nature the principle of moving in a straight line; or, in other words, toward a place where it is impossible to arrive, there being no finite end. For nature, as Aristotle well says himself, never undertakes that which cannot be done.
>
> (Galileo, EN, VII, 43)

Therefore, "the most that can be said for straight motion is that it is assigned by nature to its bodies (and their parts) whenever these are to be found outside their proper places, arranged badly, and are therefore in need of being restored to their natural state by the shortest path" (Galileo, EN, VII, 56); but in a well-arranged world only circular motion, about multiple centers, is the proper natural local motion of natural bodies. Although the Copernican physics that Galileo was reaching for would eventually be built upon the primacy of straight, not circular, motion, the Aristotelian cosmology and its underlying physics could not survive the conceptual criticism of Galileo. For, as he lets Salviati say, "whenever defects are seen in the foundations, it is reasonable to doubt everything else that is built upon them" (Galileo, EN, VII, 42). No wonder that, pace Feyerabend, Aristotelianism ceased to be, for Galileo's ablest readers, a viable intellectual option.

The most famous and perhaps also the clearest example of conceptual criticism issuing in a scientific revolution is Einstein's discussion of the classical concept of time in §1 of "Zur Elektrodynamik bewegter Körper" (Einstein 1905b). To understand him properly we must bear in mind that the kinematics of Newton's *Principia*, purportedly based on the transcendent notions of absolute space and time, gave way in the late 19th century—at least in the more enlightened circles—to the revised critical version of Newtonian kinematics proposed by Carl Neumann (1870) and perfected by James Thomson (1884) and Ludwig Lange (1885). Neumann and his followers developed the concept of an inertial frame of reference, which is Einstein's starting point. In fact, Lange's definition of an inertial frame—which, by the way, is very close to Thomson's—is much more appropriate to Einstein's needs than the one that he himself, somewhat carelessly, gives[33] and was therefore appositely prefixed by Max von Laue to his masterly exposition of Special Relativity.[34]

Lange defines an "inertial system" as a frame of reference in whose relative space three given free particles projected from a point in non-collinear directions move along straight lines. Following Neumann, Lange defines an "inertial time scale," i.e., a time coordinate function adapted to an inertial frame, by the following stipulation: A given free particle moving in the frame's space traverses equal distances in equal times, measured by the scale in question. Let F be an inertial frame endowed with an inertial time scale t. Relatively to F and t the Principle of Inertia can be stated as an empirically testable law of nature: Any free particle that is not involved in the definition of F or t travels with constant speed in a straight line.

What apparently nobody realized until Einstein made it obvious is that the Neumann-Lange definition of an inertial time does not determine a unique partition of the universe into classes of simultaneous events. If F and t are as above and x, y, and z are Cartesian coordinate functions for the relative space of F, then any real-valued function linear in t, x, y, and z,

$$t' = a_0 t + a_1 x + a_2 y + a_3 z + a_4$$

is also an inertial time scale adapted to F. (The transformation $t \mapsto t'$ rotates each hyperplane $t =$ const. about its intersection with the axis $x = y = z = 0$.)

Einstein overcame this ambiguity with his famous definition of time by means of radar signals emitted from a source at rest in the chosen inertial frame:

> If at a point A of space there is a clock, an observer at A can time the events in the immediate neighborhood of A by finding the positions of the hands that are simultaneous with these events. If there is at the space point B another clock—and we wish to add, "a clock with exactly the same constitution as the one at A"—it is possible for an observer at B to time the events in the immediate neighborhood of B. But without further stipulations it is not possible to compare, with respect to time, an event at A with an event at B. We have so far defined only an "A time" and a "B time" but no common "time" for A and B. The latter time can now be defined by stipulating *by definition* that the "time" required by light to travel from A to B equals the "time" it requires to travel from B to A. Let a ray of light start at the "A time" t_A from A towards B, let it at the "B time" t_B be reflected at B in the direction of A and arrive back at A at the "A time" t'_A. By definition the two clocks synchronize if
>
> $$t_B - t_A = t'_A - t_B$$
>
> (Einstein 1905b, pp. 893–94)

Einstein's stipulation determines a time coordinate function unique up to the choice of origin and unit, the *Einstein time* of the frame. Let t be the Einstein time of an inertial frame F. Relatively to F and t the Principle of the Constancy of the Velocity of Light can be stated as an empirically testable law of nature: Any optical signal that is not involved in the definition of t travels in vacuo with the same constant speed in a straight line, regardless of the state of motion of its source. Einstein's Principle of Relativity says that the laws of physics take the same form when referred to any kinematic system consisting of Einstein time and Cartesian space coordinates adapted to an inertial frame. The joint assertion of the Principle of Relativity and the Principle of the Constancy of the Velocity of Light entails that any two such kinematic coordinate systems are related to each other by a Poincaré transformation.[35]

All the revolutionary implications of the Special Theory of Relativity follow from this result. Let me recall one only. If t and t' are time coordinate functions defined by Einstein's method, employing the same time unit, for two inertial frames F and F' that move past each other with speed v, then the partition of nature into classes of simultaneous events determined by t is different from and incompatible with the one determined by t'. Specifically, for any event E and any arbitrary positive real number T there always exist events (or, at any rate, possible event locations) E_1 and E_2 such that $t(E) = t(E_1)$

$= t(E_2)$, but $t'(E) - t'(E_1) = t'(E_2) - t'(E) = T$. In other words, for any event E, there are events simultaneous with E by t that, by t', precede or follow E by as much time as one chooses.[36] From the inception of Special Relativity in 1905 this feature of the theory has been regarded as a radical departure from the classical conception of time. Since time enters into the definition of the basic kinematical concepts of velocity and acceleration and the latter is tied by Newton's Second Law of Motion to the key dynamic concepts of force and mass, the breach between Special Relativity and Newtonian mechanics could well be such as to make them truly incommensurable. Indeed, the conceptual differences between both theories apparently run so deep that one may even come to doubt that they are genuine alternatives. For obviously, if two pieces of scientific discourse refer by incommensurable concepts to incommensurable matters, neither of them can be offered as a substitute for the other.[37]

However, Einstein's point at the beginning of "Zur Elektrodynamik bewegter Körper," §1, is not that the classical conception of physical time is *wrong* or *inconvenient* and therefore ought to be replaced by his, but rather that classical kinematics *does not have* a definite notion of time sufficient to determine the relations of simultaneity and succession between distant events. Einstein does not give there any reason for *modifying* a given time concept but proceeds to *establish* one where none so definite and far-reaching was yet available. He takes for granted the standard notion of an inertial frame of reference endowed with Cartesian space coordinates. He notes that "if we wish to describe the motion of a particle, we give the value of its coordinates as functions of the time." But, he goes on to say, "such a mathematical description has a physical meaning only if one is quite clear as to what is here to be understood by time" (Einstein 1905b, p. 892). Implying that such clarity was missing in the extant literature of mathematical physics, Einstein then proceeds without further comment to fill this gap.

He considers three procedures for dating and timing events.[38] The first one we continually use. It consists in assigning to any given event the time shown by a clock when that event happened.

> Such a definition suffices in fact for the purpose of defining a time for the place where the clock is located; but it is not sufficient where it is a question of temporally connecting series of events which occur in different places or—what amounts to the same—of assigning temporal values to events which occur at places distant from the clock.
>
> (Einstein 1905b, p. 893)

The second procedure is the naïve extension of the first one to remote events.

It consists in assigning to them the time shown in the observer's clock when he sees them happen. This is the method I would normally use to record the time at which an aeroplane fell into the lagoon I see from my window, or a telephone call was made to me from overseas. But, as Einstein says, "we know by experience"[39] that this procedure "has the disadvantage that it is not independent of the standpoint of the observer with the clock" (Einstein 1905b, p. 893).[40] Therefore Einstein came up with the third procedure, whose description I have earlier quoted. He assumed *as a matter of physical fact* that this procedure can be consistently applied and that it does not depend on the spatial location of the base point A (nor presumably on the time t_A at which the synchronization is performed).

On these assumptions, Einstein has little difficulty in showing that, when the time variable that occurs in the equations of classical mathematical physics is understood in the manner just proposed by him, Maxwellian electrodynamics satisfies the Principle of Relativity, i.e., the Maxwell equations hold in their standard form in every inertial frame of reference if they hold in one. The experimentally recorded insensitivity of the speed of light to a substitution of frames can therefore be accounted for in a most natural way. No difference in the speed of light is *measured* when it is referred to different inertial frames because, when time is understood in Einstein's sense, the speed of a given light signal in vacuo happens to *be* the same in all such frames.

By exposing the lack of definiteness of the Newtonian time concept Einstein undermined the entire stock of notions built upon or intertwined with it. However, neither he nor his fellow physicists yielded to the cheap temptation of dismissing Newtonian science as one big connected piece of nonsense. On the contrary, they were at pains to show how, in the light of the new mode of thought, the Newtonian system possessed both meaning and truth, within appropriate limits. The equations of relativistic kinematics and mechanics collapse into Newtonian equations in the limit $(v/c)^2 \to 0$ (where c denotes the speed of light in vacuo and v the greatest speed achieved by the material objects under consideration). Consequently, *according to the new theory*, the Newtonian equations *hold good* wherever $(v/c)^2$ is negligible. But, of course, as P. M. Churchland (1979, p. 85) aptly notes, what is thereby vindicated is not the Newtonian theory as conceived by its founders but only a simulacrum of it, intellectually parasitic on Special Relativity, from which it obtains its meaning. Thus, with hindsight, we can confer definite denotations to the Newtonian concept of time and the other Newtonian concepts that depend on it; namely, the same as the homonymous relativistic concepts would have when $v \ll c$. By this move, Special Relativity inherits all the empirical evidence that was once supposed to corroborate the Newtonian theory, while the latter attains finality within its henceforth definite domain of validity.[41]

Einstein's critique of Newtonian chronometry is exceptional, for both its deadly efficacy and the pervasiveness of the concept under fire. It is no wonder, therefore, that the transition from classical to relativistic mechanics has been so often adduced as an example in discussions of the incommensurability thesis. From what we have just seen, it follows that in this particular case the thesis is adequate, but innocuous. There can be no factual basis for comparing two theories of kinematics when one possesses and the other lacks a definite criterion for timing events. But this difference alone is sufficient to give the former a crushing advantage over the latter.

In other transitions in the history of scientific thought the effects of conceptual criticism, though important, have been less decisive. Einstein (1905a) argued that the possibility of deriving the so-called Rayleigh-Jeans law of blackbody radiation from classical electrodynamics and statistical mechanics proved the inconsistency of classical physics.[42] But the successive quantum theories introduced in the 20th century to account for the emission and absorption of radiation and related phenomena have not been hitherto more satisfying from a purely intellectual point of view than the classical theories they dislodged. It is not on account of their greater conceptual perfection that the quantum theories have so far prevailed. Indeed, it is unlikely that anybody has ever thought of justifying a preference for the quantum approach to microphysics only, or mainly, because the classical theory of radiation is inconsistent.

Issuing from the very mode of thought it dissects and dissolves, conceptual criticism keeps a steady grip on whatever facts the former had got hold of while at the same time improving the way they are understood. Thus, any chasms that might arguably open between successive stages of intellectual history can be effectively bridged. Shall they, however, remain unbridged in major scientific revolutions in which conceptual criticism is secondary and indecisive? As I noted in Section 2.3, conceptual chasms need not be feared if our reference to objects and objective situations is altogether independent of the concepts by which we grasp them. To counter the claims of incommensurabilism with such a conception of reference was the aim of the theory of meaning we shall now examine.

2.6 Reference without sense

The realization that scientific thought is not referred to its proper objects through a set of easily decidable, theory-neutral, "observable" predicates was one of the main motives for the new theory of meaning developed by Hilary Putnam (1973, 1975).[43] Putnam thought he could kill the incommensurability thesis of Feyerabend and Kuhn by severing the traditional link between the reference or denotation of scientific terms and their connotation or sense. For then the former can remain stable even as the latter undergoes upheaval. Since Putnam no longer believes in reference without sense (see Section 2.6.5), I have some qualms about criticizing him on this issue. However, other philosophers still back Putnam's former theory of meaning,[44] and now they may even draw sustenance from Putnam's revival of some of his early arguments and examples in Chapter 2 of *Representation and Reality* (1988).[45] Rather than argue with them, I have addressed my polemic to their source, although it involves fighting the straw man that Putnam left behind as he moved to a more insightful philosophical position.

2.6.1 Denoting and connoting

Traditionally, a general term is said to *denote* any and every object of which it is true, and it is said to *connote* the conditions that an object meets if the term is true of it.[46] These characterizations suggest that by fixing the denotation of a term its connotation will be automatically taken care of; for it comprises the features shared by the denotata as a matter of fact. But fixing the denotation is not without difficulty, as the following examples should make clear.

Consider a binary predicate, such as 'x is heavier than y', or 'x sits on y', or 'x laughs at y'. Each of these terms is true of an ordered pair $\langle\alpha,\beta\rangle$ if a true statement is obtained by substituting a name of α for x and a name of β for y. Therefore, the denotation of a given binary predicate consists of the ordered pairs of which that predicate is true. But, what is an ordered pair? One is tempted to say that an ordered pair is a pair of things taken in a certain order, and leave it at that; for one can hardly come up with clearer words to give a more perspicuous answer. But 20th century logicians, intent on explicating everything in terms of sets—i.e., of amorphous collections identified only by their members, regardless of any order or other relations between them—produced other, presumably less naïve, definitions of an ordered pair. To my mind, the one that best renders our naïve intuitions is the following:

$$\langle\alpha,\beta\rangle =_{\text{Df}} \{\{\alpha,1\},\{\beta,2\}\} \tag{1}$$

(In other words, the ordered pair with first element α and second element β is a set of two sets, each containing two elements, viz., the object α and the number 1, and the object β and the number 2, respectively.)

But definition (1) will not satisfy someone who feels that the positive integers are not sufficiently perspicuous. Logicians have therefore adopted the ingenious definition of an ordered pair proposed by Kuratowski (1921):

$$\langle\alpha,\beta\rangle =_{\text{Df}} \{\{\alpha\},\{\alpha,\beta\}\} \tag{2}$$

Yet evidently the following would do just as well:

$$\langle\alpha,\beta\rangle =_{\text{Df}} \{\{\alpha,\beta\},\{\beta\}\} \tag{3}$$

Or, recalling that the empty set $\varnothing$ is included in every set (see Section 2.8.1), and is therefore universally available, one can use it as a marker and define an ordered pair with Wiener:

$$\langle\alpha,\beta\rangle =_{\text{Df}} \{\{\alpha\},\{\beta,\varnothing\}\} \tag{4}$$

Surely, the reader can conceive of still other alternatives.[47] Now, when one sets up the formal semantics of an artificial language one may indeed fix by decree the denotata of binary predicates by arbitrarily choosing any plausible definition of an ordered pair and excluding all others. But this solution is not open to the student of living languages. Thus, for example, there is no serious reason for maintaining that the English expression '*x* owns *y*' denotes each and every set $\{\{\alpha\},\{\alpha,\beta\}\}$ such that α owns β rather than, say, each and every set $\{\{\alpha\},\{\beta,\varnothing\}\}$ such that α owns β.

If the world is indeterministic, the very fact that all discourse is temporally and spatially localized severely restricts the determinateness of reference. General terms used today in ordinary conversation denote—unless otherwise indicated—objects that exist now or have existed in a not too distant past or are expected to exist in a more or less imminent future, on or near the Earth. 'Cow', spoken by a farmer, refers to the domestic animal whose milk we drink, not to its wild forebears, let alone to creatures biochemically indistinguishable from our cows living in a remote galaxy. But scientific discourse—at any rate in physics and chemistry—purports to be universal in scope. When we say 'electron' or 'sodium chloride', presumably we denote every electron or every molecule of sodium chloride that was, is, or will be. In a deterministic universe, the set of all electrons, past, present, and future, can be regarded at any time and place as a perfectly definite aggregate, identified by its members, to which the term 'electron' refers. But in an indeterministic world, in which electrons are being continually created and annihilated at random, no such a *set of all electrons* can—here and now—be identified by membership. It appears, therefore, that in a world like ours, any attempt to conceive the reference of general terms of universal scope as a relation in the sense of set theory, i.e., as a set of ordered pairs, is doomed not just for linguistic but for physical reasons. For suppose that the reference of a term T is understood as a binary relation whose domain—i.e., the set of the first elements of each pair—consists of all the utterances or inscriptions of T and whose codomain—i.e., the set of the second elements of each pair—consists of all the objects denoted by T. Evidently, if the identity of a set is fixed exclusively by the identity of its members, such a relation will be defined at a given time and place only if the elements of both its domain and its codomain are determinate then and there. Now, although the domain of the said relation can often meet this requirement (viz., if the term T belongs to a dead language or is no longer in use), the codomain normally cannot meet it if T is a general scientific term of universal scope and the universe is indeterministic.

The set of denotata of a general term does not give rise to the stated difficulty if its identity is given not by its actual membership but by the necessary and sufficient conditions for belonging to it; in other words, if the denotation of the term depends on its connotation. This is, of course, the traditional view, according to which the referents of a general term are picked out by the concept expressed by that term. Putnam developed his ideas on meaning in overt opposition to this view.

2.6.2 *Putnam's attack on intensions*

Putnam holds against the traditional view the fact that nobody has yet explained what it is "to grasp an intension" (Putnam, PP, vol. II, pp. 199, 263). Now, this expression, often employed by Carnap to speak of the mental operation of conceiving the connotation of a general term, is no doubt unfortunate. For it suggests that the intension of such a term, i.e., the condition or conditions that an object must satisfy in order to be denoted by it, stands in some supernatural place—such as the proverbial "museum in which the exhibits are meanings and the words are labels" (Quine 1969, p. 27)—ready to be seized with the mind's hand.[48] But if one ignores this anyway silly suggestion, it is not easy to see the force of Putnam's objection, for there are plenty of situations in life that nobody has ever explained and yet undeniably occur. To mention just one example: I cannot explain what it is for me to perceive the pencil I hold between my fingers, and indeed I cannot even imagine what it would be for me to explain it; but my inability in both these respects does not detract from the truth of the statement that there are pencils and that I am now writing with one. We might, perhaps, list the conditions—physical or otherwise—without which my perception of the pencil could not take place; but no such list could teach, say, a bodiless spirit what it is to see and to feel and to hold a pencil. The task of elucidating what it is "to grasp an intension" is still more hopeless. For surely one cannot, without circularity, explain in general the understanding of conditions by listing the conditions of understanding.

Putnam assumes for the sake of the argument that grasping the intension of a given term *T* consists in a specific, as yet unanalyzed, perhaps unanalyzable, state of mind and proceeds to show, by means of counterexamples, that in order to refer successfully to the objects denoted by *T* it is neither necessary nor sufficient to be in such a state. It is not *necessary*, for Putnam himself succeeds in referring exhaustively and exclusively to beech trees with the term 'beech' although he is not able to distinguish them from elms or to explain the difference between these two kinds of trees. Now, this example does not show that Putnam is in the same state of mind when he uses the term 'beech' as when he uses the term 'elm'; for surely in the former case he is aware of saying or writing 'beech', not 'elm', and these words do not sound or look alike. But, of course, the denotation of a word cannot be determined by its acoustic or visual shape. Indeed, I can imagine a planet exactly like ours—which I call, after Putnam, Twin Earth—in which a language is spoken, exactly like English, except that in it 'beech' denotes elms and 'elm' denotes beeches. A professor of philosophy in Twin Harvard can then be in exactly the same mental state when he refers to elm trees as our Professor Putnam is when he refers to beech trees.

The example points to a fact about language which, according to Putnam, nobody seems to have noticed before him:

> There is *division of linguistic labor.* We would hardly use such words as 'elm' and 'aluminum' if no one possessed a way of recognizing elm trees and aluminum metal; but not everyone to whom the distinction is important has to be able to make the distinction.
>
> (Putnam, PP, vol. II, p. 227)

But the example does not prove that the expert guardians of the social meaning of a given term can fix its denotation without conceiving its connotation. Suppose that Professor Putnam moves to a large estate in the south of England that he has received as a birthday present from a grateful and wealthy student. Suppose that he asks the agricultural advisor he has brought over from America to mark with a $\mathfrak{B}$ all the beech trees in the estate, so that he may in his daily rides learn to recognize this kind of tree. As he imparts this instruction Professor Putnam succeeds in referring to just the trees he wants marked, although he is himself unable to single them out, because other speakers of English can, if need be, do it for him. One of them is the advisor, who, however, can accurately fulfil his employer's wishes not because he is acquainted with each and every beech tree in the estate but, I dare say, because he has mastered what the term 'beech' connotes.

Maybe Putnam would have granted this much even in his early days.[49] He maintained, however, that no purported mastery of connotations is *sufficient* to ensure successful reference to the objects denoted by a general term. To prove it, Putnam introduced Twin Earth and invited the reader to imagine that the oceans, lakes, brooks, plants, animals, of that planet are filled with a liquid that possesses all the phenomenological properties of water as they were known ca. 1750 but which is not H_2O but some other chemical compound we may call *XYZ*. Then, according to Putnam, our word 'water' has never referred to the liquid on Twin Earth; for 'water', even as it was used in the early 1600s, e.g., in the King James Bible, denoted the chemical compound in our seas and rivers, not one that only outwardly resembles it. Nor has 'water' in Twin English ever referred to water, but only to *XYZ*, even before Twin Cavendish established its chemical composition. Therefore, assuming that science and literature have evolved in both planets more or less in the same way, we must conclude that, when Twin Shakespeare made Twin Antony say, pointing at the swiftly changing clouds,

> That which is now a horse, even with a thought
> The rack dislimns, and makes it indistinct
> As water is in water,

he could not refer to the same fluid as our Bard in the familiar homophonic lines; although both poets may have had very much the same in mind as they wrote these passages. Thus, two terms can connote exactly the same conditions to their users and yet denote different things.

Perhaps it is no accident that Putnam had to resort to this farfetched story to drive his point home. If we take him at his word, we must believe that *XYZ* will quench a man's thirst, in spite of the fact that, as J. B. S. Haldane once remarked, even the pope is 70% water.[50] Contemporary chemistry would be unable to account for such an extraordinary phenomenon. But, having consented to play Putnam's game of science fiction, we may just as well imagine that Twin Earthian scientists have in this respect been more fortunate than ours. As they do not parse matter in the manner of Lavoisier and his successors, their conception of the substance in their sea has continued to fit water also. They will not be surprised, therefore, if they find that our liquid mixes well with their blood. Imagine, moreover, that finally they have been able to explain, from their non-Lavoiserian standpoint, the hitherto baffling phenomena of electrolysis, etc. that lend support to our chemistry and its untenable distinction between H_2O and *XYZ*. In this version of the story it may still be true that 'water' in current scientific English refers to the terrestrial and not to the Twin Earthian fluid, but in scientific Twin English 'water' refers to both. It is up to the reader to decide what the two Shakespeares meant by this word in the last line of the above quotation. Maybe, with poetical insight, each referred to H_2O by one occurrence of 'water' and to *XYZ* by the other. Thus, if we let science fiction embrace not only the subject matter of science but also the manner in which scientific thought deals with it, it can teach us a lesson directly opposed to the one that Putnam tried to derive from it. Give a mild Kuhnian twist to Putnam's tale and the denotation of 'water' will turn out to depend on accepted scientific theory.[51]

2.6.3 The meaning of natural kind terms

But we do not have to indulge in fiction to perceive the inanity of Putnam's attempt to stabilize the reference of scientific terms by divorcing it from scientific thought. It is enough that we take a look at the theory of meaning proposed by Putnam himself. He developed it for two classes of general terms only, namely, for natural kind terms, such as 'elm', 'molybdenum', 'neutrino', and for physical magnitude terms, such as 'mass', 'electric charge', 'heat'. According to him the meaning of a natural kind term is fully specified by four components: (i) a *syntactic marker* indicating the grammatical classification—and behavior—of the term; (ii) a *semantic marker* indicating the ontological type or category to which the entities denoted by it belong; (iii) a *stereotype* or list of features that are normally shared by those entities and which a language user is required to know before he can be said to have mastered the term (e.g., *to be striped* is part of the English stereotype of 'tiger'—although albino tigers are not striped); and (iv) the term's *extension*. Thus, the meaning of 'water' may be given by the following ordered quadruple: (i) mass noun, concrete; (ii) natural kind,[52] liquid; (iii) colorless, transparent, tasteless, thirst-quenching, etc.; (iv) H_2O (give or take impurities). Putnam stresses that when the meaning of 'water' is explained in this way "this does *not* mean that knowledge of the fact that water is H_2O is being imputed to the individual speaker or even to the society. It means that (*we* say) the extension of the term 'water' as *they* (the speakers in question) use it is *in fact* H_2O" (Putnam, PP, vol. II, p. 269). But the extension of a natural kind term must be somehow determined, if not *by* the language users, at any rate *for* them. According to Putnam this can be achieved only if their use of the term is connected "by a certain kind of causal chain to a situation" in which the term is instantiated (Putnam, PP, vol. II, p. 200; cf. p. 176). The extension of the term is thereby fixed; it comprises anything that is *the same* as the said instance, in the manner of sameness indicated by the semantic marker of the term.[53]

> Our theory can be summarized as saying that words like 'water' have an unnoticed indexical component: 'water' is stuff that bears a certain similarity relation to the water *around here.* Water at another time and in another place or even in another possible world has to bear the relation same_L [i.e., *same liquid as*] to *our* 'water' *in order to be water.*
>
> (Putnam, PP, vol. II, p. 234)

In this theory of meaning, despite Putnam's protestations,[54] the extension of a natural kind term still depends on its intension in two respects. In the first place, the particular instance of the term—e.g., the water around here—to which its denotation is anchored must be conceived under the category—stuff, liquid, chemical compound, or whatever—designated by the term's semantic marker. Users of the term must therefore understand that it refers, e.g., to a liquid body and not, say, to its surface or its glitter, etc. In the second place, the conditions for being *the same* as the said instance with regard to that category—e.g., for being the same liquid or the same chemical compound as the water around here, or the same sort of tree as the elms over Abbie Cabot's roof—must be fixed, lest the reference of the term be unsteady. I do not contend that very many natural kind terms in common use meet this last requirement but only that their denotation is well defined just to the extent that they do meet it. Now, the conditions for being the same as a given object with respect to a certain appropriate category constitute precisely what one would normally call the connotation or intension of a natural kind term. And surely it is not too bold to claim that both the categories distinguished by the semantic markers of our language and the conditions of sameness with regard to them fall under the jurisdiction of philosophical and scientific thought and are liable to change with it.

2.6.4 Speaking of quantities

Putnam treats physical magnitude terms more cursorily than natural kind terms. This is a pity, for magnitudes are central to modern science, and their manner of being—and of being signified—is of great interest to philosophy. Putnam's approach to the matter is, if I may say so, one of studied laxity. He labels a section of one of his semantical papers "The Meaning of Physical Magnitude Terms" (Putnam, PP, vol. II, pp. 198–207), and he tells us—in a different paper—that it contains "a causal theory of reference in connection with magnitude terms" (Putnam, PP, vol. II, p. 176n.). But, except for the remark that the users of such terms know that they refer to "putative physical *quantities*— capable of more and less, and capable of location" (Putnam, PP, vol. II, p. 199), we do not find in the said section any considerations about meaning that would not apply equally well to natural kind terms. Yet surely speaking of quantities implies some commitments that are not presupposed when one talks about natural kinds.

It is not clear to me whether Putnam considers physical magnitudes a species of physical quantities or whether he regards 'quantity' and 'magnitude' as synonymous. Aristotle, who first discussed the category of quantity, said that a 'quantum' (πόσον) is a 'multitude' (πλῆθος) if it can be counted, a 'magnitude' (μέγεθος) if it can be measured. In the same passage, Aristotle defined a quantum as "that which is divisible into constituents each of which is by nature a *one* and a *this*." He also noted that multitudes can be divided into discrete parts, whereas magnitudes are potentially divisible into constituents that are mutually connected or continuous (συνεχῆ) (Aristotle, *Metaphysica*, Δ, 13, 1020^{a}7–11). Thus, for Aristotle, a 'magnitude' was a quantity of a certain sort, viz., continuous quantity, the only sort that can be measured.

Of course, modern physics has scored great successes in measuring properties like temperature or angular momentum that are not divisible into constituents each of which is a *one* and a *this*.[55] Therefore, we no longer subscribe to Aristotle's definitions but allow the category of quantity to encompass any attributes of things, processes, or events that "can be reasonably represented numerically" (Krantz et al. 1971, p. xvii). In an interesting essay on "The Mathematical Classification of Physical Quantities," J. C. Maxwell attributes to Hamilton the "most important distinction" between "Scalar quantities, which are completely represented by one numerical quantity, and Vectors, which require three numerical quantities to define them" (Maxwell 1890, vol. II, p. 259).[56] For example, Newtonian mass is a scalar quantity and Newtonian force a vector quantity. And there is, of course, no reason why we should not also speak of tensor quantities, like stress. As to scalar quantities, some—like probability—are real-valued (i.e., representable by real numbers); while others—like the amplitude of a quantum-mechanical probability wave—are complex-valued. Does Putnam's term 'magnitude' refer to all these different sorts of quantities? An answer to this question cannot be extracted from his text. I shall assume, however, that Putnam's 'magnitudes' are real-valued scalar quantities. This will greatly simplify our discussion without restricting its scope, for the main conclusion we shall reach regarding such quantities holds a fortiori for the rest.

An attribute of physical objects is a magnitude in the stated sense if, but only if, its particular instances can be assigned numbers in such a way that some of the distinctive structural features of these numbers come to represent physical relations between those instances. For example, our ordinary concept of mass requires that the mass of a body *A* should be assigned a number equal to *n* times that ascribed to the mass of another body *B* whenever *A* balances a body formed by putting together *n* copies of *B*. (The copies in question need only be equal *in mass*, i.e., they must balance one another.) The investigation of the exact import and requirements of such numerical representations of physical attributes was initiated by Helmholtz (1887) and carried forward by Hölder (1901) and N. R. Campbell (1920). It reached a high level of sophistication in the treatise *Foundations of Measurement*, by Krantz, Luce, Suppes, and Tversky (1971). Further advances are recorded in Louis Narens' *Abstract Measurement Theory* (1985). In the light of this research tradition, the numerical representation of an attribute of physical objects is best conceived as a mapping of the set of its instances into the field **R** of real numbers, so contrived that some of the structural features of **R** faithfully represent—by way of the homonymous mappings induced on Cartesian products and power sets—relations characteristic of that set. The existence of such a mapping obviously entails that the particular instances of the attribute provide a concrete realization of a species of structure—in Bourbaki's sense—whose distinguished components meet specific conditions. (The terminology employed and the general mathematical ideas involved in the two foregoing sentences and in the rest of this section are explained in the appendix on mathematical structures, which is Section 2.8. Readers who are bored by mathematical definitions may proceed at once to the conclusion in the last paragraph of the present subsection, on page 65, provided that they are willing to accept it on faith.)

To give an idea of the nature and strength of such conditions I shall report on the two simplest types of extensive magnitude characterized in *Foundations of Measurement.* Roughly speaking, extensive magnitudes—like mass—are numerically representable attributes whose instances can be compared as to size and can be added to one another to make larger instances. The numerical representation of such magnitudes makes use of two structural features of the real number field, namely, its additive group structure and the ordering of the real numbers by the binary relation '≥'. As a matter of fact, what is used is not the full additive group structure of **R** but the semigroup structure that remains when addition is restricted to the set $\mathbf{R}^+$ of the positive reals, while their inverses and the neutral element 0 are forgotten. It is thus apparent that in order to conceive a physical attribute as an extensive magnitude one must distinguish at least two structural features in the extension Q of that attribute, viz., (i) a reflexive, connected and transitive relation P such that $\langle Q,P\rangle$ is a weak order, and (ii) an associative binary operation $\oplus$ such that $\langle Q,\oplus\rangle$ is a semigroup. ($\oplus$ is therefore a mapping of Q^2 into Q. We write $a\oplus b$ for the value of $\oplus$ at $\langle a,b\rangle$. $\oplus$ is associative, so that for any $a,b,c\in Q$, $(a\oplus b)\oplus c = a\oplus(b\oplus c)$.) A faithful numerical representation of $\langle Q,P,\oplus\rangle$ will then be given by a mapping μ: $Q\rightarrow\mathbf{R}^+$ such that for all $a,b\in Q$,

M1. aPb if and only if $\mu(a)\geq\mu(b)$; and

M2. $\mu(a\oplus b)=\mu(a)+\mu(b)$.

Note that if μ satisfies M1 and M2, and $\alpha\in\mathbf{R}^+$, $\alpha\mu$: $a\mapsto\alpha\mu(a)$ is a mapping of Q into $\mathbf{R}^+$ which also satisfies M1 and M2 and thus provides another faithful numerical representation of the quantity Q. μ and $\alpha\mu$ are said to be different *scales* for measuring the same quantity.

A mapping μ meeting conditions M1 and M2 exists if and only if $\langle Q,P,\oplus\rangle$ is what Krantz et al. call a closed positive extensive structure, i.e., if and only if its components satisfy the axioms C1–C5 given below.[57] For greater perspicuity, I write '$a\sim b$' for 'aPb and bPa'; and 'na' for '$a_1\oplus\ldots\oplus a_n$', where $a_i\sim a$ for every value of the index i $(1\leq i\leq n)$.

$\langle Q,P,\oplus\rangle$ is a *closed positive extensive structure* if and only if

C1. $\langle Q,P\rangle$ is a weak order, and for all $a,b,c\in Q$,

C2. $(a\oplus b)Pa$, but it is not the case that $aP(a\oplus b)$;

C3. $(a\oplus b)\oplus c\sim a\oplus(b\oplus c)$;

C4. The following three conditions hold whenever any one of them holds: aPb, $(a\oplus c)P(b\oplus c)$, and $(c\oplus a)P(c\oplus b)$;

C5. If aPb but it is not the case that bPa, then for any $x, y \in Q$, there is a positive integer n such that $(na \oplus x)P(nb \oplus y)$.

Axioms C1–C5 are consistent if the theory of real numbers is consistent, for the structure $\langle \mathbf{R}^+, \geq, + \rangle$ evidently would satisfy them.

One may well doubt that the instances of any extensive physical magnitude can in fact be weakly ordered by size. If they could, the relation 'x is neither greater nor smaller than y' (here symbolized by '$\sim$') would be transitive. Evidently, this requirement is not satisfied by observed magnitudes, for due to the limited power of resolution of our instruments and organs, any three instances a, b, and c of an extensive magnitude can be such that the size of a is indistinguishable from that of b, and the size of b from that of c, and yet a is perceptibly larger than c. However, it is ordinarily assumed that the actual instances of an extensive magnitude that our observations do but partially and imperfectly record are weakly ordered by size (just as, mutatis mutandis, the actual instances of an intensive physical magnitude, such as temperature, are understood to be weakly ordered by intensity).[58] On this understanding, if P designates the ordering relation, every extensive magnitude that is studied in physics satisfies axiom C1. Moreover, the instances of such a magnitude can, under certain conditions, be joined in a standard way to make larger instances. Evidently, if $\oplus$ designates the operation of joining two instances of a given magnitude in the appropriate way, axiom C2 is fulfilled by any instances of it, a and b, which can be thus joined. The reader will be persuaded that C4 is satisfied if c is an instance that can be joined to a and to b, and that C3 holds good if c can be joined to $a \oplus b$. Now, two instances of an extensive magnitude cannot be joined to make a bigger one if any of them happens to be the product of joining the other with a third one. But this limitation is easily overcome by treating instances of equal size as interchangeable copies. Then, even if $a = b \oplus c$, $a \oplus b$ is defined if there is an instance $b' \sim b$, and a can be joined to b'. The solution fails, though, if we run out of copies. A similar—and equally insuperable—limitation arises if one of the instances is too large to be joined with the other by the appropriate standard method. To cope with this difficulty, one can include in the structure of any extensive magnitude Q for which it arises a distinguished subset D of Q^2, on which alone the operation of joining is defined, and treat $\oplus$ as a mapping of D into Q. A faithful numerical representation of such a structure $\langle Q, P, D, \oplus \rangle$ is given by a map μ: $Q \rightarrow \mathbf{R}^+$ that meets condition M1, but satisfies the following requirement instead of M2:

M2*. For every $a, b \in D$, $\mu(a \oplus b) = \mu(a) + \mu(b)$.

Krantz et al. have established sufficient conditions for the existence of a mapping μ satisfying M1 and M2*. These conditions involve, of course, the replacement of C2, C3, and C4 by other axioms that make allowance for the fact that the operation $\oplus$ is now defined only on a part of Q^2. They also require a drastic revision of axiom C5, which we shall now discuss.

Axiom C5 is a version of the so-called Archimedean postulate. It says in effect that if the set Q of instances of an extensive magnitude is the base set of a closed positive extensive structure, no two instances of that magnitude are incommensurable. No matter how small is x and how large is y, if a is but slightly larger than b, there is an integer n such that x joined to n copies of a is larger than y joined to n copies of b. But if $\oplus$ is not everywhere defined on Q^2, it may not be possible to form na and nb for a sufficiently large integer n. Yet the faithful representation of an extensive magnitude in the semigroup $\langle \mathbf{R}^+,+\rangle$ certainly demands that the structure of that magnitude should satisfy some form of postulate. The version chosen by Krantz et al. to overcome the present difficulty can be simply stated if we introduce two new terms. Let *a standard sequence in Q* designate either an infinite sequence $a_1, a_2, \ldots$ or an m-tuple $\langle a_1, \ldots, a_m\rangle$ of elements of Q, such that for each positive integer n less than the number of terms in the standard sequence, $a_{n+1} = a_n \oplus a_1$. Let us say that a standard sequence in Q is *strictly bounded* if there is a b in Q such that for every term a_k in the sequence, bPa_k is true but a_kPb is not. The Archimedean postulate for $\langle Q,P,D, \oplus\rangle$ says then that *every strictly bounded standard sequence in Q is finite* (i.e. is an m-tuple, for some positive integer m). In effect, Krantz et al. (1971, p. 87) prove that there is a mapping μ: $Q \rightarrow \mathbf{R}^+$ satisfying M1 and M2* if $\langle Q,P,D, \oplus\rangle$ is what I shall call an *ordinary extensive structure*, that is to say, if

E1. $\langle Q,P\rangle$ is a weak order, and for all $a,b,c \in Q$,

E2. If $\langle a,b\rangle \in D$, then $(a \oplus b)Pa$, but it is not the case that $aP(a \oplus b)$;

E3. If $\langle a,b\rangle \in D$ and $\langle a \oplus b,c\rangle \in D$, then $\langle b,c\rangle \in D$, and $\langle a,b \oplus c\rangle \in D$, and $((a \oplus b) \oplus c)P(a \oplus (b \oplus c))$;

E4. If $\langle a,c\rangle \in D$ and aPb, then $\langle c,b\rangle \in D$ and $(a \oplus c)P(c \oplus b)$;

E5. If aPb but it is not the case that bPa, then there is some $x \in Q$ such that $\langle b,x\rangle \in D$ and $aP(b \oplus x)$;

E6. Every strictly bounded standard sequence in Q is finite.

Krantz and his associates characterize further structure for periodic extensive magnitudes (e.g., angle), extensive magnitudes with an essential maximum,[59] and many types of nonextensive magnitudes. Additional structures are defined by Narens (1985). But the two examples examined here sufficiently illustrate the complexity and far-reaching scope of some of the structural conditions involved in quantitative thinking. To use a physical magnitude term correctly one certainly does not have to know the Krantz-Luce-Suppes-Tversky axioms for the corresponding structure, any more than one needs to know the Hilbert axioms for the Euclidian plane in order to pave the kitchen floor with square tiles. But just as a commitment to thus pave a floor presupposes that the Hilbert axioms apply to it within the admissible margin of imprecision, so a meaningful reference to a physical attribute as an extensive magnitude of a certain type implies that its instances stand to one another as elements of the pertinent structure. Indeed even the humblest thought of an extensive quantity—e.g., the thought that regulates the pouring of oil into a food processor to make mayonnaise—conceives its instances as orderable by size and as liable to increase by the addition of further instances. Such notions, plus an idea of the standard method or methods for joining its instances, belong to what, after Putnam, we may call the stereotype of a given extensive magnitude term. But, of course, in a Putnamist theory of meaning, the truly critical factor in the semantics of a general term is not its stereotype but its extension. The extension of a natural kind term is the set of its instances. Insofar as this set is open-ended, it can only be determined by a list of conditions (an intension). The same can be said of the instances of a magnitude, but with an added proviso: a collection of particulars can pose as the set of instances of a certain magnitude only to the extent that it is a realization of the corresponding abstract structure. While the extension of a natural kind term is normally conceived as an unstructured set, a set can be the extension of a given physical magnitude only if it is the base set of a suitable structure. Of course, no collection of particulars actually met in science is seen as *the* set of (all conceivable) instances of a magnitude. But in order to be grasped as *a* set of such instances it must be embedded in the full structure.[60] A fortiori, the same holds also for nonscalar physical quantities.

It would be interesting to inquire into the epistemological implications of this remarkable difference between the thought of quantities, prevalent in modern science, and the thought of natural kinds, favored by Aristotle. One surmises that Aristotle's neglect of quantitative concepts and the tightly knit structures they connote is the main reason why, notwithstanding his conviction that natural necessity is the distinguishing mark of the subject matter of natural science, he did so poorly at unravelling any particular examples of it. But here we must deal with the meaning of physical magnitude terms, and specifically with the likelihood that they preserve their reference when they are inherited from a scientific theory by its revolutionary successor. Putnam contends that a physical magnitude term securely enjoys a stable reference when its uses have the proper causal connection with a situation—an "introducing event"—in which the magnitude referred to by that term was actually singled out "as the physical magnitude *responsible* for certain effects in a certain way" (Putnam, PP, vol. II, p. 200). Even if the introducing event has been forgotten, the intention of referring to the same magnitude that was referred to by the same term in the past links our current use of it to those earlier uses. Indeed, the very presence of the term in our vocabulary is a causal product of earlier events and ultimately of the introducing event.

> If anyone knows that 'electricity' is the name of a physical quantity, and his use of the word is connected by the sort of causal chain I described before to an introducing event in which the causal description given was, in fact, a causal description of electricity, then we have a clear basis for saying that he uses the word to refer to electricity. Even if the causal description failed to describe electricity, if there is good reason to treat it as a mis-description *of electricity* (rather than as a description of nothing at all)—for example, if electricity was described as the physical magnitude with such-and-such properties which is responsible for such-and-such effects, where in fact electricity is responsible for the effects in question, and the speaker intended to refer to the magnitude responsible for those effects, but mistakenly added the incorrect information 'electricity has such-and-such properties' because he mistakenly thought that the magnitude responsible for those effects had those further properties—we still have a basis for saying that both the original speaker and the person to whom he teaches the word use the word to refer to electricity.
>
> (Putnam, PP, vol. II, p. 201)

I take it that when Putnam speaks of physical quantities being *responsible* for observed effects he does not mean to anthropomorphize them, but only to describe them as efficient causes. Now, causality has been conceived in philosophy as a relation between two things (or rather, between a person or, generally, an animal and a thing, as when Aristotle said that the sculptor was the cause of the statue) or between two events (as in 'the fire was caused by lightning'). Perhaps common sense would also acknowledge causal connections between things and events (as in 'a single fission bomb caused the death of thousands of Japanese children'). But I confess I had not heard of causal relations between magnitudes and phenomena. I must therefore wait for a new analysis of causality before I can accept the causal efficacy of magnitudes *as such,* not just of the things or events that sport them. Indeed, even if such an analysis were forthcoming, I doubt it could succeed in bestowing causal powers on magnitudes like distance, or time, or vertical acceleration. Putnam wisely chooses not to talk about them but takes as an example a physical magnitude term that was long used as the name for a putative physical substance. Some of this obsolete connotation must still attach to the current stereotype of 'electricity' if, as Putnam says, it includes the idea of "a magnitude which can move or flow." Such thinglike quantities are met perhaps at "introducing events," but for a straight physical magnitude term which is not thus categorially ambiguous it is very hard to imagine wherein such an event might consist. For my part, I do not think that if I had stood "next to Ben Franklin as he performed his famous experiment" and had heard him say that 'electricity' denotes that "which collects in clouds" until it suddenly "flows from the cloud to the earth in the form of a lightning bolt," I could have even guessed that he was introducing a name for a *magnitude,* not a fluid.

Be this as it may, the important lesson that can be drawn from our discussion of extensive magnitudes is that he who conceives a particular as an instance of a quantity must place it wittingly or unwittingly in the relational network of a structure constituted by it and its fellow instances. If a quantity is an attribute that can be represented numerically, its conditions of identity must anyway include the structural features portrayed by its numerical representations. The reference of a physical magnitude term cannot be impervious to changes in the structure that keeps its extension together.

2.6.5 'Mass' in classical and relativistic dynamics

The foregoing considerations lend support to the incommensurabilist claim that the physical referent of the term 'mass' in relativity physics is by no means identical with that of the homonymous Newtonian term (Kuhn 1962, p. 101). Feyerabend (1962) stated this claim very simply, as follows:

> In classical, pre-relativistic physics the concept of mass [. . .] was absolute in the sense that the mass of a system was not influenced (except perhaps causally) by its motion in the coordinate system chosen. Within relativity, however, mass has become a relational concept whose specification is incomplete without indication of the coordinate system to which spatiotemporal descriptions are all to be referred.
>
> (Feyerabend 1981, vol. I, p. 81)

The coordinate-dependent quantity to which Feyerabend refers here is presumably what is known as *relativistic mass* ('apparent mass' for Dixon 1978, p. 114), i.e., the scalar factor by which one must multiply the velocity of a particle to obtain its relativistic momentum. Since the relativistic mass of a particle P is the same with respect to all coordinate systems relative to which P has the same speed, its numerical representation may be regarded as defined on the set of ordered pairs $\langle P,v\rangle$, where v stands for the particle's speed and ranges over the closed-open interval $[0,c)$. This alone constitutes an unbridgeable difference between relativistic mass and the quantity called 'mass' in classical mechanics.[61]

However, besides relativistic mass, Relativity assigns to every material particle a scalar quantity known as *proper mass* (also 'rest mass'), which, like classical mass, is a function of state, i.e., is independent of the particle's motion. This bifurcation of the meaning of 'mass' makes it at first blush even harder to equate the term's reference in relativistic and in classical dynamics (see Field 1973).[62] But, on a closer examination, it will be apparent that the proper mass has a better right than the relativistic mass to take over the role of the classical concept of mass. Indeed, a very striking analogy can be drawn between the laws of motion of classical and relativistic dynamics when they are stated in an idiom that makes them comparable. In the standard four-dimensional formulation of Special Relativity developed by Minkowski (1908, 1909), the following equation holds:

$$\mathrm{F} = \dot{\mathrm{p}} = m\ddot{\mathrm{r}} \tag{5}$$

Here r is a four-vector representing the worldpoint or spacetime location of a particle (referred to an arbitrarily chosen origin in Minkowski spacetime); m, p, and F denote, respectively, the particle's proper mass, its four-momentum and the four-force acting on it; and a dot over a variable signifies differentiation with respect to the particle's proper time. Now, not only is eqn. (5) typographically almost indistinguishable from Newton's Second Law of Motion in its standard three-dimensional formulation (I have refrained, however, from using boldface for the four-vectors), but it is homologous with the corresponding law of classical mechanics in the neo-Newtonian four-dimensional formulation (as explained, e.g., by Friedman 1983). To obtain a statement of this law we do not have to rewrite equation (5); it is enough to reinterpret it by letting m, p, and F stand, respectively, for the particle's classical mass, its neo-Newtonian four-momentum, and the neo-Newtonian four-force on it, and by taking the dot over a variable to mean differentiation with respect to universal time. Even the latter seemingly drastic reinterpretation loses much of its sting if we recall that in a neo-

Newtonian theory universal time (i.e., the time coordinate into which all Einstein times collapse in the limit $c \rightarrow \infty$) agrees with each particle's proper time along its particular worldline.

There is, however, one very substantial difference between classical mass and the proper mass of Relativity: the proper mass of a particle changes in inelastic collisions (collisions in which the aggregate kinetic energy of the colliding particles is not conserved). Now, this is not one of those differences between two theories that concern only our beliefs about some entity but not the meaning of the term denoting it. For inelastic collisions are of course inevitable in any process of putting masses together to make larger masses. In fact, according to Relativity, when several particles are brought together, the aggregate's proper mass includes not only the mass equivalent of the kinetic energy lost as their relative velocities vanish but also the mass equivalent of any work done against forces that tend to keep the particles apart (*minus,* of course, the work that would have to be done, to separate the particles, against their binding forces). In sharp contrast with this, the total classical mass of a system of classical particles is simply the sum of the masses of each, no matter what their nature and our manner of joining them. Since the standard physical methods of adding the instances of a physical extensive quantity play a key role in the constitution of its structure, the relativistic term 'proper mass' cannot share the same reference with the classical term 'mass' even if they occur in typographically identical and conceptually kindred equations. This may not be altogether obvious as we draw worldlines on paper, labelling each with a little number equal to its proper mass, while at the same time forgetting that those lines stand for material bodies which exert forces on each other and harbor an internal structure which will be set in commotion whenever the lines touch. But as soon as one is reminded of it, it becomes clear that if proper mass has an extensive structure, it must be a very peculiar one, quite different from the straightforward ordinary extensive structure of classical mass.[63]

It is amusing to fancy what an "introducing event" for 'mass' would look like according to the causal theory of reference. It evidently will not do to have an imaginary Newton tell his assistant, after toiling in vain to move a large stone, "Look, it's the stone's big *mass* that won't let it budge," for if the stone were placed on top of a wheelbarrow he would readily carry it away, barrow and all. However, as writers of fiction we are entitled to move the scene to the age of railways. Our Newton could then point out to his hearers that in order to come to a full stop a loaded train needs, ceteris paribus, a much longer piece of track than an empty train. "*Mass,*" he might then proclaim, "is the physical quantity responsible for this effect." The trouble is that if the fancy world he lives in happens to be one in which there is an upper bound to the speed of signals, the quantity thus designated—like the track length used for estimating it—would be frame-dependent. In other words, our mock-Newton would have introduced a magnitude akin to relativistic mass, not to classical and proper mass. To have him introduce by ostension the quantity *m* in equation (5) we must place him with his students in a more recherché situation. For instance, they may travel together in a squadron of freely falling spaceships of different sizes. He can then declare *mass* to be that which causes each ship to abide by its geodesic worldline, so that the larger the mass, the more fuel you must burn to make it deviate from free fall. Obviously, to catch their master's meaning, the students must do more than just look and hear: they must think hard and recall a good deal of the differential geometry and the geometric theory of gravity they learned before taking off. There is no such thing as thoughtless ostension, but in cases like this one the intellectual sophistication involved in ostensively introducing a term is apt to be considerable. As a matter of fact, in real life 'mass' was not introduced by any of the methods suggested here. It was in the course of theological discussions concerning the Holy Eucharist that Newton's term for classical mass, viz., 'quantitas materiæ', came to signify the quantity that remains unchanged when the volume of a body changes by condensation and rarefaction.[64]

2.6.6 Putnam's progress

After 1976, Putnam by and large disassociated himself from the views on reference we have examined in this section. He announced his new stance in his presidential address to the American Philosophical Association on December 29, 1976 (Putnam 1978), and gave a careful statement of its grounds and implications exactly one year later, in his presidential address to the Association for Symbolic Logic (Putnam 1980). Here he distinguished three main positions on reference and truth: "the extreme Platonist position, which posits non-natural mental powers of directly 'grasping' forms," "the verificationist position which replaces the classical notion of truth with the notion of verification or proof, at least when it comes to describing how the language is understood," and "the moderate realist position which seeks to preserve the centrality of the classical notions of truth and reference without postulating non-natural mental powers" (Putnam 1980, p. 464; also in Putnam, PP, vol. III, pp. 1–2). He then proceeded to show, on the strength of the Löwenheim-Skolem Theorem and related results of the theory of models for first-order formal languages, that his earlier "moderate realism"—i.e., realism without "non-natural" grasp of intensions—must be given up. This is not the place to go through Putnam's model-theoretic argument.[65] The gist of it is that the "moderate realist," who tries to carry on with scientific discourse, à la Putnam 1975, by means of general terms which denote without connoting, is incapable of singling out the "intended" interpretation of any scientific theory he may put forward (in a first-order language), from among the infinitely many distinct and even incompatible interpretations that satisfy it.

> Nor do 'causal theories of reference', etc., help. Basically, trying to get out of this predicament by *these* means is hoping that the *world* will pick one definite extension for each of our terms even if *we* cannot. But the world does not pick models or interpret languages. *We* interpret our languages or nothing does.
>
> (Putnam 1980, p. 482; PP, vol. III, p. 24)

> If 'refers' can be defined in terms of some causal predicate or predicates in the metalanguage of our theory, then, since each model of the object language extends in an obvious way to a corresponding model of the metalanguage, it will turn out that, *in each model M, reference*$_{\mathrm{M}}$ is definable in terms of *causes*$_{\mathrm{M}}$ but unless the word 'causes' [. . .] is already glued to one definite relation with metaphysical glue, this does not fix a determinate extension for 'refers' at all.
>
> (Putnam 1980, p. 477; PP, vol. III, p. 18)

As Putnam will have nothing to do with preternatural Platonistic insights into ready-made mind-independent intensions, he seeks a way out of his predicament in the example of mathematical intuitionism and constructivism.

> 'Objects' in constructive mathematics are *given through descriptions.* Those descriptions do not have to be mysteriously attached to those objects by some nonnatural process (or by metaphysical glue). Rather the possibility of *proving* that a certain construction (the 'sense', so to speak, of the description of the model) has certain constructive properties is what is asserted and *all* that is asserted by saying the model 'exists'. In short, *reference is given through sense, and sense is given through verification procedures and not through truth conditions.* The 'gap' between our theory and the 'objects' simply disappears—or, rather, it never appears in the first place.
>
> (Putnam 1980, p. 479; PP, vol. III, p. 21)

Putnam extends this approach to the entire philosophy of science. He disparages his former stance as "the perspective of metaphysical realism," for which "there is exactly one true and complete description of 'the way the world is'," and "truth involves some sort of correspondence relation between words or thought-signs and external things and sets of things." To this "externalist perspective" and its preferred "God's Eye point of view" he opposes "the internalist perspective," thus called "because it is characteristic of this view to hold that *what objects does the world consist of?* is a question that it only makes sense to ask *within* a theory or description" (Putnam 1981, p. 49).

> In an internalist view also, signs do not intrinsically correspond to objects, independently of how those signs are employed and by whom. But a sign that is actually employed in a particular way by a particular community of users can correspond to particular objects *within the conceptual scheme of those users.* 'Objects' do not exist independently of conceptual schemes. *We* cut up the world into objects when we introduce one or another scheme of description. Since the objects *and* the signs are alike *internal* to the scheme of description, it is possible to say what matches what.
>
> (Putnam 1981, p. 52)

> What is wrong with the notion of objects existing "independently" of conceptual schemes is that there are no standards for the use of even the logical notions apart from conceptual choices. [. . .] We can and should insist that some facts are there to be discovered and not legislated by us. But this is something to be said when one has adopted a way of speaking, a language, a "conceptual scheme." To talk of "facts" without specifying the language to be used is to talk of nothing; the word "fact" no more has its use fixed by the world itself than does the word "exist" or the word "object."
>
> (Putnam 1988, p. 114)

Those of us who woke early from dogmatic slumber can hardly be surprised by Putnam's "internalism," but we certainly welcome it. Its relevance to the Feyerabend-Kuhn incommensurability thesis and related problems of scientific change is clear enough.[66] Its implications, however, can only be gauged if the notion of a conceptual scheme is made more precise.

2.7 Conceptual schemes

While the phrase 'conceptual scheme' is a fairly new addition to the philosopher's stock-in-trade—there is still no entry for it in the index to Paul Edwards' *Encyclopedia of Philosophy* (1967)—the thoughts it is meant to express or suggest can be traced back to Kant's "Copernican revolution" in philosophy. Foremost among them is the distinction between the sensuous content or "matter" of empirical knowledge and its rational ordering or "form," i.e., the universal outline or *scheme* that is being progressively filled by the manifold "given" through observation and experiment.[67] This distinction leads almost at once to the idea that "the same" content might be structured—by extraterrestrials? by our children's children?—according to a different scheme. Kant was emphatic that "we cannot render in any way conceivable or comprehensible to ourselves" alternative "forms" of sense awareness and discursive thought, "even if they should be possible," and that "even assuming that we could do so, they would still not belong to experience, the only kind of knowledge in which objects are given to us" (Kant 1781, p. 230). But the thought of liberating one's fellow men from the bonds of their inherited mode of thinking and experiencing and of achieving immortality by conjuring up a newly shaped world has proved irresistibly attractive and it still rings titillatingly in many a comment on the revolutionary scope of the Theory of Relativity or the shattering implications of Quantum Mechanics. One further connotation of 'conceptual scheme' can be read out of the above quotation from Kant. Besides being formal or *schematic* and admitting alternatives (within or beyond our reach), a conceptual scheme is *systematic*—it weaves experience into one coherent whole, leaving no loose ends or poorly fitted links a different scheme could seize upon in an attempt to supplant it. This opinion of the wholesomeness of our "forms" of perceiving and understanding is not surprising in a philosopher who, for all his critical radicalism, continued to view them—in the manner of 18th century creationism—as springing from the God-given nature of human reason (see Kant 1781, p. 669); but from our present evolutionary perspective it looks indeed most implausible. Perhaps that is why nowadays the completeness of "our conceptual scheme" and its capacity to digest anything we might come across are not often argued for. And yet, if it were openly acknowledged that our variegated experience is being put together in accordance with diverse, imperfectly coherent, partly asystematic, unfinished, not altogether stable patterns of understanding, much recent philosophizing about conceptual schemes—pro or contra—would lose its edge.[68]

The two writers who have probably done most to disseminate the English expression 'conceptual scheme' on either side of the Atlantic are W. O. Quine and Sir Peter Strawson. The Kantian themes of form versus content and of all-embracing order, as well as the post-Kantian theme of alternative, disposable schemes, resound in the following passage of Quine's "On What There Is," published in 1948:

> We adopt, at least insofar as we are reasonable, the simplest conceptual scheme into which the disordered fragments of raw experience can be fitted and arranged. Our ontology is determined once we have fixed upon the over-all conceptual scheme which is to accommodate science in the broadest sense; and the considerations which determine a reasonable construction of any part of that conceptual scheme, for example, the biological or the physical part, are not different in kind from the considerations which determine a reasonable construction of the whole.
>
> (Quine 1961, pp. 16–17)

Here Quine appears too sanguine about our chances of finding a vantage point outside all conceptual schemes from which to judge them and make an intelligent choice among them. Later in life, he withdrew this suggestion, noting that the study and revision of a given conceptual scheme cannot be undertaken "without having some conceptual scheme, whether the same or another[. . .], in which to work" (Quine 1960, p. 276). The principles of simplicity and coherence that preside over conceptual reform are now said to be internal to the prevailing scheme.

In his profound and influential book *Individuals: An Essay in Descriptive Metaphysics* (1959), Strawson voices a far more moderate opinion than Quine's about the scope available for conceptual change:

> Certainly concepts do change, and not only, though mainly, on the specialist periphery; and even specialist changes react on ordinary thinking. [But] there is a massive central core of human thinking which has no history—or none recorded in histories of thought; there are categories and concepts which, in their most fundamental character, change not at all. Obviously these are not the specialties of the most refined thinking. They are the commonplaces of the least refined thinking; and are yet the indispensable core of the conceptual equipment of the most sophisticated human beings. It is with these, their interconnexions, and the structure that they form, that a descriptive metaphysics will be primarily concerned.
>
> (Strawson 1959, p. 10)

Throughout the book, Strawson refers to this unchanging "core of human thinking" by the phrase "our conceptual scheme."

Neither Quine nor Strawson nor—as far as I know—any of their followers have attempted to give a full, systematic description of such a scheme, in the manner of Kant's "metaphysic of experience," or of the many systems of categories that 19th century philosophers developed in Kant's wake. The cautiousness of our contemporaries stems, no doubt, in part from a deep seated aversion to heavy-handedness in philosophy; but it may also be due to some however dim realization that no such system is to be had. Be that as it may, what seems to be agreed on by all writers who countenance this general approach is that our current adult way of singling out, identifying, and reidentifying particular objects of diverse sorts pertains to our conceptual scheme. This fits well with what Putnam says in the passage quoted near the end of Section 2.6. Not all conceptual schemers, however, will accept Putnam's claim that, because we "cut up the world into objects when we introduce one or another scheme of description," it follows that objects "do not exist independently of conceptual schemes." A more subtle view has been firmly and clearly put forward by David Wiggins:

> [Philosophy] must hold a nice balance [. . .] between the extent to which the concepts that we bring to bear to distinguish, articulate and individuate things in nature are something invented by us and the extent to which these concepts are something we discover and permit nature itself to intimate to us and inform and regulate for us. Conceptualism properly conceived must not entail that before we got for ourselves these concepts, their extensions did not exist autonomously, i.e. independently of whether or not the concepts were destined to be fashioned and their compliants to be discovered. What conceptualism entails is only that, although horse, leaves, sun and stars are not inventions or artifacts, still, in order to single out these things, we have to deploy upon experience a conceptual scheme which has itself been fashioned or formed in such a way as to make it *possible* to single them out.
>
> (Wiggins 1980, p. 139)

Wiggins' balancing act is not without difficulty, but it is preferable—at any rate, for and within our human self-understanding—to Putnam's facile plunge into the quagmire of relativism. For, as Wiggins aptly emphasizes, to single out something, one must single *it* out. Reference must reach out for the referent, not bring it about, or it will have failed its purpose. Positing objects and pointing at them—by word or gesture—are both necessary, equally respectable, but altogether different mental functions. A philosophy that is unable to distinguish between them lacks, so to speak, sufficient power of resolution. The myth of the self-differentiating object ("the object which announces itself as the very object it is to the mind"—Wiggins 1980, p. 139) is in effect a remnant of primeval animism and we ought to disown it. But it is then all the more inevitable that we own the self-differentiating *subject* as the living principle of all experience and, indeed, of all truth. Now, self-differentiation presupposes that *from* which that which I call myself differentiates itself. It also imposes constraints on any conceptual scheme that might conceivably be my own. Let us take a somewhat closer look at this.

The method of "transcendental deduction," especially as practiced by Fichte, was mainly directed at spelling out such constraints. It was assumed that self-awareness in general (*Selbstbewußtsein überhaupt*) is possible, and it was purportedly shown that its conditions of possibility include all the main features of our physical and cultural world, such as space, time, causation, other minds, even private property. Such derivations normally suffered, in one way or another, from the same vice, viz., that a specific content was surreptitiously imported, at one or more stages of the argument, from our human experience into the generic conditions presupposed by the initial assumption. By artfully scaffolding several such steps at each of which a little more was taken for granted than had been validly proved by the preceding ones, the transcendental philosopher managed to go all the way from the abstract demand that a wholly unspecified form of self-awareness must be possible right down to the familiar structures of everyday life. The fallacies of transcendental deduction can be avoided by restricting the inquiry to our human self-awareness and the constraints on conceptual schemes that are manifestly inherent in it. This, if I am right, was Strawson's intent in *Individuals.* Like all genuine philosophy, an inquiry along these lines involves a considerable risk of error. But it is not condemned, like the argument from *Selbstbewußtsein überhaupt,* to a choice between sophistry and vacuity. Of course, it will not disclose the "only possible" conceptual scheme that "consciousness in general" can wield. But its findings may nevertheless be sufficient to curb relativism without yielding to superstition.

We are aware of ourselves as persons surrounded by things and other persons with whom we interact physically. These conditions of our human self-awareness do not entail that we may not wake up one day and find ourselves as disembodied spirits amidst a choir of angels—whatever that may mean. But even granting, for the sake of the argument, the unlikely premise that our own identity would not perish in such a sweeping change, it is apparent that the identity of our world cannot survive it. Now, a "form of experience" to which no transition can be made from ours without loss of our self-identity or of the identity of "the real" from which we self-differentiate ourselves is plainly irrelevant to the philosophy of human knowledge. Thus, self-differentiation, as we know it from our self-awareness, restricts the variety of conceptual schemes that, in any interesting sense, deserve our consideration.

It may well be that the idea of a conceptual scheme which is inaccessible from ours, and yet is exercised on the same contents, is not absurd in God's eye. But a God's eye viewpoint is clearly one that we cannot share. Besides, although we can make good sense of the abstract thought of a scheme of categories or basic modes of conception and predication apart from any empirical contents, nobody has ever been able to attach a definite meaning to the notion of a pure content of experience apart from a categorial framework; yet this notion is presupposed by the idea of mutually inaccessible or, as the saying goes, incommensurable schemes of thought that have the same field of application in God's eye.

In point of fact, none of the historical examples adduced by the incommensurabilists—neither the dislodgement of epicycle astronomy by Kepler or phlogiston chemistry by Lavoisier or ether electrodynamics by Einstein, nor even the displacement of Aristotelian by mathematical physics—has brought about an immediate, total remaking of the European mind. In particular, these turnabouts in science do not appear to have affected the enduring core of thought mentioned by Strawson in the passage quoted above. This comprises the conceptual means employed in identifying and reidentifying, localizing and dating—i.e., generally speaking, in objectifying—the diverse features of our everyday life. The stability of such a conspicuous part of our intellectual resources has no doubt provided some of the motivation for the philosophical teachings of Kant and logical positivism about unsurpassable limits to conceptual innovation, discussed in Section 2.4. Since deep and far-reaching conceptual changes are known to have happened in science, a critical point in all such doctrines is the relationship they establish between the inert core of everyday thinking and what Strawson superciliously describes as "the specialist periphery."

On one interpretation of Kant, the connection he saw here was simple enough: the basic concepts and principles of classical mathematical physics were only an improved, streamlined version of the scheme by which men had mentally organized their environment from time immemorial. This view, however, is not only questionable in itself (what precedent is there in everyday thought for the classical idea of the evolution of a physical system governed by differential equations?); it becomes utterly preposterous when one tries to extend it to the connection between the "permanent core" and 20th century physics.

The philosophy of logical positivism was developed in full consciousness of the volatility of fundamental physics, and its mature version keeps clear of such difficulties. No system of scientific thought has an intrinsic, privileged connection with human reason (as Newtonianism supposedly had according to Kant). The ordinary and the "specialist" concepts are mutually related by arbitrary "correspondence rules." Scientific theories may be invented and elaborated in blissful ignorance of actual experience. However, their application to it rests on the suitable choice of a partial interpretation of their esoteric terms in plain words. That such application might involve—and require—a new way of conceiving experience is out of the question. The aim of scientific theorizing is prediction, not understanding. Things are well understood as they are commonly understood.

Scientific change would certainly be less disquieting if all it had to offer were new, more successful ways of telling where the pointers will stop in certain dials given that they stand within such-and-such intervals in other dials. A dial, however, owes its specific difference to the instrument bearing it, which in turn, apart from the scientific ideas that presided over its construction, is just a meaningless "black box." Hence scientific thought must supply the interpretation of ordinary laboratory talk—explaining, for instance, what this or that pointer reading says—and not the other way around. But thanks to modern science and technology we now live among boxes which, for all their gleaming screens, are closed and "black" to the common understanding. With them the "specialist periphery" has penetrated our daily lives and undermined the "core of human thinking." Not that we need to understand their workings in order to use them for our ends. But they stand on every side bearing witness to the inadequacy of our prescientific conceptual scheme for properly grasping what we can see and handle.

It is a commonplace of contemporary culture that the differences between prescientific and scientific categories preclude a satisfactory description in ordinary terms of many of the commonest artifacts of science. We ought to realize, moreover, that there would have been no room for such differences to arise or for the insufficiency of the traditional scheme of thought to become manifest, had that scheme been the definite, comprehensive, coherent, and inherently stable system that Kant and other like-minded "descriptive metaphysicians" made it out to be. If our categorial framework were complete and closed in itself, it would leave no gaps through which to catch a glimpse of anything beyond it, and our understanding could grow only by the further specification of existing ideas, not by the invention of new ones. Fortunately, however, our so-called conceptual scheme is not the ready-made blueprint for a True Intellectual System of the Universe but rather a motley of patterns of understanding, to which others can still be added by varying, extending, or simply forgetting those already available.

The flexibility of our understanding is well demonstrated by the variety of criteria we bring to bear on the identification of the ordinary objects of our attention. Some philosophers maintain that we grasp all such objects as substances or as attributes—properties or relations—of substances. But the very difficulty they have in reaching a coherent and satisfactory account of substances should make us wary of oversimplification in this matter.[69] Suppose you are asked to hum a tune you have just heard played on the piano. Physics has taught us to understand a tune as a complex pattern of air waves; but this understanding is quite irrelevant to your task. Either the category of substance must be stretched to encompass musical notes and melodies, or you have here to do with an object which you grasp neither as a substance nor as an attribute. A flash of lightning, a gust of wind, a fire, a river, can, with some effort of the imagination, be brought under the substance-and-attribute scheme. But it is not under any such description that they are thought of when they ordinarily attract our attention. A waterfall is falling water, but it is individuated by the falling, not by the water. A philosopher might argue that it is a feature—an "accident"—of the place where it is located; but in ordinary experience it is the fall that makes the place, ordering the entire landscape around the stupendous downpouring of ever-renewed water. It is permissible and perhaps even plausible to hold with Aristotle that all happenings like those just mentioned are either motions, or alterations, or shrinkings and growths, or generations and destructions, of substances. But this is a principle of revisionary, not descriptive, metaphysics; a guide for the formation of scientific hypotheses, not a truth about the actual structure of prephilosophical human experience.

The inadequacy of the substance-and-attribute scheme for capturing our ordinary ways of thinking is especially obvious if one seeks to apply it to social and cultural realities. Naturalistically minded ontologists have labored in vain to reduce to it the familiar concepts and concept clusters under which those realities are effortlessly grasped, e.g., 'art', 'language', 'money', 'law'. What substances change as a language gradually loses the subjunctive or when a loan to a developing country is classified as "nonperforming"?

But even within its accustomed range of application the classical category of substance is too heterogeneous to play the unifying role usually assigned to it. Aristotle declared that "substance is thought to belong most obviously to bodies" and went on to explain that we therefore say that "animals and plants and their parts are substances," but also "the natural bodies, such as fire and water and earth and the like, and such things as are either part of them or composed from them, [. . .] like heaven [οὐρανός] and its parts, stars and moon and sun" (*Metaphysica*, Z, 2, 1028^{b}8ff.). Water and "earth," however, are not usually grasped as vast, disconnected, sprawling bodies, but rather as stuffs out of which bodies are made. (The same, presumably, must be said of fire, as Aristotle understood it; although one should be permitted to doubt that flames were ever perceived as substantive bodies except by philosophers and very small children.) The classical prototypes of substance are thus quite disparate: on the one hand, the "simple bodies" or stuffs; the living organisms, on the other. The tension between these two extremes was often resolved by allowing only one of them and discarding the other. While the godforsaken majority tended to equate primordial being or substance with stuff—a preference already implicit in the use of the word *substantia,* which latinizes the Greek ὑποκείμενον, not οὐσία—a bold and imaginative thinker like Leibniz chose the opposite option: according to him, it is from my self-awareness and my awareness of other beings like myself that I obtain my idea of substance.[70] Of course, Aristotle embraced stuffs and plants and animals and even his incorporeal astrokinetic intelligences in the category of substance, and much of his lasting influence may be due to this broad-mindedness. But for him it is "a man or a plant or some other thing such as these that we pronounce to be substances above anything else" (ἃ δὴ μάλιστα λέγομεν οὐσίας εἶναι—*Metaphysica,* Z, 7, 1032^{a}19). And in his doctrine of the simple bodies—earth, water, air, fire, and ether—each endowed with its own characteristic internal principle of motion and rest which is like a simple appetite, he somehow conceives their nature on the analogy of an organism's life. Aristotle laid thereby the foundations of an admirably coherent and comprehensive system of the world. But he also exposed himself to criticism such as that reported above (Section 2.5) and drove the study of motion into a blind alley.

Galileo and his successors, who in the 17th century established a new science of motion, turned away from substances and their natures and sought instead the laws of phenomena. They renounced Aristotle's encyclopedic ambitions and devoted their attention to special types of events displayed in well-circumscribed situations. They subjected them to a new way of thinking that is, to this day, the intellectual hub of physics. It consists in developing for any given such type of occurrence a *physical theory,* which in effect sustains the unity of the type and accounts for its internal diversity. In Chapter 3, we shall examine the more significant features of this mode of thought in the light of some recent work in the philosophy of science. Before proceeding to it, let me summarize our findings about the question raised in Section 2.3 regarding the continuity of scientific thought and the comparability of scientific claims in the face of radical conceptual innovation. I stated there three conditions, each one of which would be sufficient to warrant such comparability and continuity, viz.:

C1. Some concepts are immune to change, and they provide a stable reference to decisive facts.

C2. The new concepts are arrived at through internal criticism of the old, by virtue of which the facts purportedly referred to by the earlier mode of thought are effectively dissolved.

C3. Reference to facts does not depend on the concepts by which they are grasped.

In Section 2.6 I argued at length that reference to facts is not independent of the concepts by which they are grasped. In Section 2.4 I contended that there are no known grounds for believing that any concepts relevant to science are immune to scientific change. In Section 2.5 I tried to show how new concepts arrived at through internal criticism of the old preserve the continuity of scientific thought, but I noted that it is unlikely that all new scientific concepts are formed in this manner. Thus, the results of our inquest are quite contrary to C1 and C3, and only partly supportive of C2.

In view of such findings, it is fair to conclude that the incommensurabilists would be right *if* our understanding were a tightly knit system in which a change in any fundamental concept must bring about semantic displacements in all the rest. In that case, revolutions in basic science would shake and shatter the very roots of reference, and successive, conceptually diverse scientific theories could not, strictly speaking, be about the *same* objective situations or compete for confirmation by the *same* empirical evidence. However, our human reason is not the rigid, single-purpose, all-of-a-piece engine of war fancied by some philosophers but is a many-faced, makeshift bundle of intellective endeavors. Indeed, Kant himself, who set so much store by the purportedly systematic nature of our categorial framework, was aware that our scientific understanding cannot cope with our moral life and our experience of art, and proclaimed the separateness of these distinct spheres of reason.[71]

Here, indeed, we are not concerned with the all-too-obvious multiplicity of our modes of thought in diverse fields and walks of life, but with intellectual variety in science and specifically in physics. The modern understanding of natural phenomena by means of physical theories methodically segregates the domain of each new theory from the broad background of experience, as articulated by common sense and earlier science. To have a limited scope—even when the limits are not exactly known—is therefore a characteristic of all physical theories (and an important reason for their effectiveness). No physical theory lays claim to a global understanding of reality.

The physicist who substitutes one theory for another goes on living in the same neighborhood, working at the same institute, driving every day the same old road between them, to and fro, while reflecting on the politics of his nation or the moods of his teenage children or the failings of his car in the light of the same social, moral, or low-level mechanical concepts as before. Kuhn's dictum that "after a revolution scientists are responding to a different world" is either a piece of empty rhetoric or evidence of a misunderstanding of the nature and scope of physical theories. The founders of mathematical physics did of course abandon the Aristotelian system, which was indeed a worldview, not a physical theory. But that does not mean that they left the world in which they and their forebears had lived until then. For it was not an Aristotelian world, and nobody, not even Aristotle himself, had ever managed to see everything in it as depicted by the Aristotelian view. Nor do theory dislodgments such as physics has repeatedly known since the 17th century imply in and by themselves a change of worldview, for neither the dislodged nor the dislodging theories have ever concerned the *world* or entailed a definite conception of it.[72]

Kuhn's dictum is indeed true—and trivial—if by "a different world" he just meant the peculiar domain of the revolutionary theory. But even on this more temperate interpretation the phrase is apt to be misleading. The domain of a revolutionary physical theory is normally conceived so as to include that of the earlier theory which it is meant to surpass. The earlier theory's success within its own domain is then accounted for by the novel theory, which thereby draws the limits of the old one. Thus, new theories in mathematical physics typically do not dislodge their predecessors except to lodge them permanently in the appropriate epistemic niche (Rohrlich and Hardin 1983). The innovative physicist must indeed be able to refer to the old domain in order to rethink it as a proper part of the new one. But this is hardly surprising, unless we picture scientists, as in the caricature Kuhn has drawn of them, on the analogy of religious zealots, who may "convert" from one mode of thinking to another but cannot retain two together in their one-track minds. In the real world, of course, the creators of new physical theories have been trained in the extant ones and—as noted in Section 2.5—it is often by critically reflecting on them that they find a way of going beyond them. All special fields of inquiry must, moreover, be accessible from the same general background of human life. Thus, for example, a relativist and an ether theorist could both work with interferometers made to the same specifications by the same manufacturer or read about Michelson and Morley's experiment in the same issue of *The Philosopher's Magazine*. We do not hold the shortsighted opinion that physical theories have no further aim than that of enmeshing select parts or features of the common background in a web of calculations. But that is not to deny that the background lies there, ill-understood, confusing even, like a murky ocean joining the shiny islands of theory.

Summing up: The "reality" from which we differentiate ourselves as thinking agents is articulated by the concepts that shape our thought and regulate our action. They provide the intellectual means by which particular objects are distinguished from one another, identified and reidentified. Radical conceptual change would therefore wreak havoc on objectivity, as the incommensurabilists maintain. Their strictures, however, cannot impair the continuous, coherent development of mathematical physics, because this is not a succession of comprehensive, mutually incompatible views of that "reality" (a succession of *Weltanschauungen*) but a plurality of interrelated attempts at conceiving definite features or aspects or parts of it by means of intellectual systems of limited scope ("physical theories," in the peculiar sense to be explored in Chapter 3).

In spite of the often deep conceptual differences between such systems, they are protected against incommensurability by the following factors:

(i) They all belong to a connected historical tradition, in which new formations normally grow from the older ones by critical reflection about them and with explicit reference to them.

(ii) They all draw many of their concepts from the same fairly coherent body of mathematical thought.

(iii) They pick out their respective domains from the one background variously known as "reality," "human experience," or, more pompously, "the world."

(iv) There is an inevitable fuzziness in the way each domain is inserted or "embedded" in the background—this, in turn, favors the gross prima facie identification of some objects referred to by diverse theories, even if the latter conceive them very differently.

2.8 Appendix: Mathematical structures

Mathematical structures came up in our discussion of physical quantities in Section 2.6.4 and will be central to our consideration of physical theories in Chapter 3. I give here a repertory of terms and symbols I use for speaking of such structures. It can serve as a refresher to readers who have some acquaintance with the subject and also as a means for controlling my terminology.

[Following the advice of one of the book's referees, I have interspersed the abstract exposition with examples, which I hope will make it accessible and useful also to readers who know very little about modern mathematics. The examples can be readily spotted because, like this paragraph, they are enclosed in brackets.]

2.8.1 Sets

I take the standpoint of a more or less naïve set theory (see Section 2.8.7). The expression '$\{a, b, c\}$' denotes the set whose members or elements are a, b, and c. If '$S(x)$' stands for a sentence in which all occurrences of a noun and all pronouns proxying for it have been replaced by the variable x, the expression '$\{x \mid S(x)\}$' denotes the set of all objects x such that $S(x)$, provided that there is such a set.

[In standard mathematics it is assumed that, given any set, the existence of certain other sets related to it is assured. Two such conditional existential assumptions are mentioned at the beginning of Section 2.8.3. It is also assumed that there is a set corresponding to the expression '$\{x \mid x$ is a natural number$\}$'. On the other hand, consider the expression '$\{x \mid x$ is a set and x is not a member of $x\}$'. Evidently, there cannot be a set which this expression denotes, for if there were such a set, it would both be and not be a member of itself.]

If a is an element of a set A, we say that a is in A, that a belongs to A, that a is contained in A, or that $a \in A$. A set's identity is completely determined by the elements it contains. If every element of a set A is also an element of the set B, we say that A is included in B, that A is a part of B, that A is a subset of B, or that $A \subset B$. Note that according to this definition, if A is any set, $A \subset A$. If $A \subset B$, the set of all elements of B that do not belong to A is called the *complement* of A in B and is denoted by $B \setminus A$.

The *intersection* $A \cap B$ of sets A and B is the set of all elements that belong to both A and B. The *union* $A \cup B$ of sets A and B is the set of all elements that belong to A, or to B, or to both. If $\mathbb{S}$ is a set of sets, the intersection of $\mathbb{S}$, denoted by $\bigcap\mathbb{S}$, is the set of all elements that belong to every set in $\mathbb{S}$; the union of $\mathbb{S}$, denoted by $\bigcup\mathbb{S}$, is the set of all elements that belong to at least one set in $\mathbb{S}$.

If A is a set, the set of all its subsets, $\{X \mid X \subset A\}$, is called the *power set* of A, and is denoted by $\mathcal{P}(A)$. We write $\mathcal{P}^2(A)$ for $\mathcal{P}(\mathcal{P}(A))$ and generally $\mathcal{P}^{n+1}(A)$ for $\mathcal{P}(\mathcal{P}^n(A))$ $(n \geq 2)$. If A and B are sets, the set formed by all ordered pairs[73] the first term of which belongs to A and the second term of which belongs to B—in other words, the set $\{\langle a,b\rangle \mid a \in A \,\&\, b \in B\}$—is called the *Cartesian product* of A by B and is denoted by $A \times B$. We write A^2 for $A \times A$, and generally A^{n+1} for $A^n \times A$ $(n \geq 2)$. A^n may be called, for brevity, the nth Cartesian product of A.

The *null* set or *empty* set $\emptyset$ is the set $\{x \mid x \neq x\}$; in other words, $\emptyset$ is the set of all objects x such that x is not identical with itself. Evidently, no object at all belongs to $\emptyset$. Thus, if A is a set, the following statement is true: Any object that happens to be an element $\emptyset$ is also an element of A. Therefore, $\emptyset \subset A$. Hence, according to the (standard) definitions adopted here, $\emptyset$ is a subset of every set. A set is said to be *non-empty* if it is not identical with $\emptyset$. Two sets A and B are said to be *disjoint* if $A \cap B = \emptyset$.

2.8.2 *Mappings*

A mapping f from a set A to a set B assigns to each element a of A one and only one element $f(a)$ of B. We often refer to such a mapping as 'the mapping $f: A \to B$ which maps A into B by $a \mapsto f(a)$'. We say that f sends a to $f(a)$. A is the *domain* of f, B its *codomain*. Every a in A is an *argument* of f; $f(a)$ is the *value* of f at a. If C is a subset of A, the mapping $f|_C: C \to B$, which maps C into B by $x \mapsto f(x)$, is called the *restriction* of f to C. If A is a subset of D, and $\mathsf{g}: D \to B$ is a mapping such that $\mathsf{g}(a) = f(a)$ for any $a \in A$, g is said to agree with f on A and to be an *extension* of f. Note that f is then the restriction of g to A. Given two mappings, $f: A \to B$ and $\mathsf{g}: B \to C$, the *composition* of f by g is the mapping $\mathsf{g} \circ f$ that maps A into C by $a \mapsto \mathsf{g}(f(a))$.

The set $\{x \mid x = f(y) \text{ for some } y \in A\}$ is a subset of the codomain of f known as the *range* of f. If the range of f is equal to the codomain of f, one says that f is a *surjective* mapping or a *surjection*, and that it maps its domain *onto* its codomain. By the *fiber* of f over b I mean the set $\{x \mid f(x) = b\}$. If every non-empty fiber of f is a singleton (i.e., if it contains a single element of A), f is said to be an *injective* mapping or an *injection*. A mapping that is both surjective and injective is said to be a *bijective* mapping or a *bijection*. If f is bijective, there is an *inverse* mapping $f^{-1}: B \to A$ such that for each $a \in A$, $f^{-1}(f(a)) =$

a. f^{-1} is plainly a bijection. Two sets are *equinumerous* if and only if there is a bijection from one to the other.

An important instance of surjection is furnished by the projection mappings attached to any Cartesian product. If $A = B_1 \times \ldots \times B_n$ is the Cartesian product of component sets $B_1, \ldots, B_n$, then A is mapped surjectively onto its kth component ($1 \leq k \leq n$) by the *projection* π_k: $\langle x_1, \ldots, x_k, \ldots, x_n \rangle \mapsto x_k$.

The *graph* of f is the set $G_f = \{\langle a, f(a) \rangle \mid a \in A\}$. Bourbaki identifies the mapping $f: A \to B$ by $a \mapsto f(a)$ with the ordered triple $\langle G_f, A, B \rangle$ (Bourbaki 1970, E.II.13). To the working mathematician this identification must seem frightfully artificial. However, by providing for the definition of a mapping in set-theoretical terms, it paves the way for the fairly simple universal characterization of mathematical structures given below.

[For example, suppose that Allison is loved by Peter and Jack, and Anne is loved by Edward and Joseph, while Mary is not loved by anybody. Let G be the set of the three girls, and B the set of the four boys. Let us also denote by A the subset of G formed by the girls whose names begin with the letter A and by J the subset of B formed by the boys whose names begin with the letter J. There is a mapping $f: B \to G$, which assigns to each boy in B the girl he loves. The fiber of f over Anne is the set {Edward, Joseph}. The restriction of f to J is an injection. There is also a mapping $\mathsf{g}: B \to A$, which also assigns to each boy in B the girl he loves. g is however a different mapping, because, although it has the same domain as f, it has a different codomain. Note that g is a surjection. Its restriction to J is a bijection. The inverse of this bijection assigns to each girl in its domain the boy in J who loves her. If we equate the ordered pair $\langle a,b \rangle$ with the set $\{\{a\},\{a,b\}\}$ (see Section 2.6.1), the graph of g turns out to be the set

{{{Peter},{Peter,Allison}},{{Jack},{Jack,Allison}},
{{Edward},{Edward,Anne}},{{Joseph},{Joseph,Anne}}}.

The reader is invited to write down the *set* g, in accordance with Bourbaki's definition of a mapping as a set. (Hint: View ordered triples as ordered pairs whose first term is an ordered pair, viz., $\langle a,b,c \rangle = \langle \langle a,b \rangle, c \rangle$.)]

2.8.3 Echelon sets over a collection of sets

In set theory the following two assumptions are normally taken for granted:

(i) If a set A is given, the power set $\mathcal{P}(A)$ is also given.

(ii) If sets A and B are given, the Cartesian product $A \times B$ is also given.

Thus, if a set of sets $\mathfrak{S}$ is given, then, by repeated application of conditions (i) and (ii) we can specify an endless array of sets that are supposedly given together with $\mathfrak{S}$. We refer to it as the array of *echelon sets* over $\mathfrak{S}$. A set A is an *echelon set* over the set of sets $\mathfrak{S}$ if and only if A meets one of the following conditions:

(α) A is one of the sets in $\mathfrak{S}$;

(β) $A = \mathcal{P}(B)$, and B is an echelon set over $\mathfrak{S}$; or

(γ) $A = B \times C$, and both B and C are echelon sets over $\mathfrak{S}$.

If the set $\mathfrak{S}$ is just the singleton $\{A\}$, we speak of echelon sets over A.

Suppose now that, after the Polish fashion, we write '$\mathcal{P}A$' for '$\mathcal{P}(A)$' and '$\times AB$' for '$A \times B$'. Then, if every set in the set $\mathfrak{S}$ is assigned a unique and exclusive name, each echelon set over $\mathfrak{S}$ can be described without ambiguity by a unique expression formed with the names of one or more sets in $\mathfrak{S}$ and one or more occurrences of the symbols $\mathcal{P}$ and $\times$. Let A be an echelon set over the set of sets $\mathfrak{S} = \{S_1, \ldots, S_n\}$ and let $Q(A)$ denote the unique expression that describes A in this way. The expression formed by substituting in $Q(A)$ the integer i for every occurrence of the name of the set S_i $(1 \leq i \leq n)$ may be called *the scheme for the echelon construction of A.* Two echelon sets over different sets of sets are said to be homologous if the schemes for their echelon construction are identical. Obviously, if the sets $\mathfrak{S}$ and $\mathfrak{S}'$ are equinumerous, then to each echelon set over $\mathfrak{S}$ there is one and only one homologous echelon set over $\mathfrak{S}'$.[74]

Consider now two equinumerous sets of non-empty sets, $\mathfrak{S} = \{S_1, \ldots, S_n\}$ and $\mathfrak{S}' = \{S'_1, \ldots, S'_n\}$. Let $f: \bigcup\mathfrak{S} \rightarrow \bigcup\mathfrak{S}'$ be such that, for every k $(1 \leq k \leq n)$, $f(S_k) = S'_k$. f determines a mapping from each echelon set A over $\mathfrak{S}$ into the homologous echelon set A' over $\mathfrak{S}'$. I call it the *homonymous mapping induced by f in A*, and when there is no danger of confusion I also denote it by f. The homonymous mapping induced by f in any echelon set over $\mathfrak{S}$ is readily defined in terms of the three conditions (α), (β), and (γ) involved in the definition of echelon sets.

(α) Let A be one of the sets in $\mathfrak{S}$. The homonymous mapping induced by f in A is the restriction of f to A.

(β) Let A and A' be homologous echelon sets over $\mathfrak{S}$ and $\mathfrak{S}'$, respectively. The homonymous mapping induced by f in $\mathcal{P}(A)$ sends each set $X \subset A$ to the set $f(X) = \{y \mid y = f(x) \text{ for some } x \in X\}$.

(γ) Let A and A', B and B', be two pairs of homologous echelon sets over $\mathfrak{S}$ and $\mathfrak{S}'$. The homonymous mapping induced by f in $A \times B$ sends each pair $\langle a,b \rangle$ in $A \times B$ to $\langle f(a), f(b) \rangle$ in $A' \times B'$.

Note, in particular, that if $V = ((U_1 \times U_2) \times \ldots \times U_n)$, where $U_1, \ldots, U_n$ are any echelon sets over $\mathbb{S}$, and f_i denotes the homonymous mapping induced by f in U_i $(1 \leq i \leq n)$, the homonymous mapping induced by f in V sends each ordered n-tuple $\langle u_1, \ldots, u_n \rangle$ in V to the n-tuple $\langle f_1(u_1), \ldots, f_n(u_n) \rangle$, which clearly belongs to the echelon set homologous to V over the codomain of f.

[Consider again the mapping f from the set B of boys to the set G of girls defined at the end of Section 2.8.2. Figure out the values of the homonymous mapping induced by f in the power set $\mathcal{P}(B)$. For instance, $f(J)$ = {Allison, Anne}, whereas f({Jack, Peter}) = {Allison}.]

2.8.4 Structures

I shall characterize a mathematical structure as a finite list or ordered n-tuple of components meeting requirements of a certain kind. The characterization could be extended to the case of a countably infinite list of components, but I shall not attempt it here.

By an *(m,n)-list of structural components* I mean an ordered n-tuple $\langle S_1, \ldots, S_m, Q_1, \ldots, Q_{n-m} \rangle$ such that (i) the first m terms $S_1, \ldots, S_m$ $(m < n)$ are distinct non-empty sets, and (ii) each one of the remaining $n - m$ terms $Q_1, \ldots, Q_{n-m}$ is a distinguished element of some echelon set over $\{S_1, \ldots, S_m\}$. The terms of class (i) are known as the *base sets*, and I shall often refer to those of class (ii) as the *distinguished components* of the (m,n)-list.

Two (m,n)-lists of structural components will be said to be *similar* if the kth term of the one is homologous to the kth term of the other $(1 \leq k \leq n)$. Let $\mathcal{L} = \langle S_1, \ldots, S_m, Q_1, \ldots, Q_{n-m} \rangle$ and $\mathcal{L}' = \langle S'_1, \ldots, S'_m, Q'_1, \ldots, Q'_{n-m} \rangle$ be two similar (m,n)–lists of structural components and put $\mathbb{S} = \{S_1, \ldots, S_m\}$ and $\mathbb{S}' = \{S'_1, \ldots, S'_m\}$. Let $f\colon \bigcup\mathbb{S} \to \bigcup\mathbb{S}'$ be a bijection such that for every k $(1 \leq k \leq m)$, $f(S_k) = S'_k$, and let f_r denote the homonymous mapping induced by f on the echelon set over $\mathbb{S}$ containing Q_r $(1 \leq r \leq n-m)$. If, for every such index r, $f_r(Q_r) = Q'_r$, I shall say that f *transports* $\mathcal{L}$ to $\mathcal{L}'$ and that $\mathcal{L}'$ is related to $\mathcal{L}$ by the *transport mapping* f. A set of conditions C jointly satisfied by the terms of an (m,n)–list of structural components $\mathcal{L}$ is *transportable* if the terms of any other such list related to $\mathcal{L}$ by any transport mapping f jointly satisfy the conditions obtained by replacing in C every occurrence of the name of each term τ of $\mathcal{L}$ by the name of $f(\tau)$.

A *species of structure* is an (m,n)–list of structural components whose terms jointly satisfy a transportable set of conditions. If $m = 1$, one usually designates any particular instance of the species of structure by the name of its sole base set.[75]

[In what follows, I shall characterize, by way of illustration, some of the more important species of structure in mathematics. However, before looking into their definitions, some readers might wish to try a hand at describing as an instance of a species of structure the situation involving two groups of boys and girls introduced at the end of Section 2.8.2. That situation involved a set B of 4 objects, a set G of 3 objects, and a mapping $f: B \to G$. One need not refer to the set of natural numbers in order to specify that G has just 3 objects. It is enough to demand that there are objects x, y, and z belonging to G such that $x \neq y \neq z \neq x$, and that any object w belonging to G is identical with either x, y, or z. This condition is clearly transportable. In a similar fashion one can say that B has just 4 objects. However, in order to define f by means of transportable conditions we need some further structure. I propose the following: Prescribe linear orders for B and G (as defined by L1–L4 below). Let $<_B$ and $<_G$ symbolize, respectively, the linear order on B and on G. I define f as a mapping which sends the first two elements of $\langle B,<_B\rangle$ to the first element of $\langle G,<_G\rangle$, and the last two elements of $\langle B,<_B\rangle$ to the second element of $\langle G,<_G\rangle$. The situation described at the end of Section 2.8.2 will be conceived as an instance of the species of structure $\langle B, G, <_B, <_G, f\rangle$ meeting the said conditions if the linear order on the set of girls is understood to agree with the alphabetical order of their names, and the linear order on the set of boys is understood to agree with the reversed alphabetical order of their reversed names.]

Let us now characterize some species of structure that are mentioned elsewhere in this book. A *group* is a quadruple $\langle G,e,f,g\rangle$, where G is an arbitrary non-empty set, e is an element of G, f is a mapping of G into G, and g is a mapping of G^2 into G such that for any elements a, b, and c of G, the following characteristic conditions are fulfilled:

G1. $g(a,g(b,c)) = g(g(a,b),c)$ (g is associative).

G2. $g(a,e) = g(e,a) = a$.

G3. $g(a,f(a)) = g(f(a),a) = e$.

g is called the group product; f, the inversion mapping; e, the neutral element.[76] It can be shown that e is unique: no further element of G can satisfy the conditions imposed by G2 and G3 on e. The group is said to be *abelian* if the product is commutative, i.e., if $g(a,b) = g(b,a)$ for all a and b in G.

A pair $\langle G,g\rangle$ meeting condition G1 is said to be a *semigroup*. A triple $\langle G,e,g\rangle$ satisfying G1 and G2 is called a *monoid*.

[For example, let G be the set of letter strings defined as follows: (i) if 'α' denotes a letter of the English alphabet, α is a letter string; (ii) if α and β are letter strings, their concatenation $\alpha\beta$ (i.e., α followed immediately by β) is a letter string. Then, if g is the mapping that assigns to any pair of letter strings α and β their concatenation $\alpha\beta$, $\langle G,g\rangle$ is a semigroup. Likewise, if G is the set of natural numbers, $\{0, 1, 2, \ldots\}$, g stands for addition on G and e denotes the number 0, $\langle G,e,g\rangle$ is a monoid. If G is the set of the positive and negative rational numbers, g stands for multiplication, f stands for the operation taking the reciprocal value, and $e = 1$, $\langle G,e,f,g\rangle$ is the *multiplicative group of rationals*. If G is the set of all rationals, g stands for addition, f stands for multiplication by -1, and $e = 0$, $\langle G,e,f,g\rangle$ is another realization of the species of structure *group*, namely, the *additive group of rationals*. There is an unending variety of groups. I shall now propose two simple examples. Let $\langle G,e,f,g\rangle$ be any group, and let f' and g' respectively denote the restrictions of f and g to $\{e\}$. Obviously, $f(e) = e$ and $g(e,e) = e$. Therefore, $\langle\{e\},e,f',g'\rangle$ is a group. Now consider an equilateral triangle ABC with center of gravity P, and let α, β, and γ denote clockwise rotations of the triangle about P by 120, 240, and 360 degrees, respectively. Note that γ agrees then with the identity mapping, which sends each point of the triangle to itself. Put $G = \{\alpha,\beta,\gamma\}$, $f(\alpha) = \beta$, $f(\beta) = \alpha$, $f(\gamma) = \gamma$, and $\gamma = e$. Let g associate with each ordered pair $\langle\xi,\eta\rangle$ of rotations in G the composite rotation $\xi\eta$ effected by applying ξ after η. Clearly, then, $g(\alpha\beta) = g(\beta\alpha) = \gamma$, $g(\alpha\alpha) = \beta$, $g(\beta\beta) = \alpha$, and, for any rotation $\xi \in G$, $g(\gamma\xi) = g(\xi\gamma) = \xi$. Thus, $\langle G,e,f,g\rangle$ is a group.]

A *linear order* is a pair $\langle S,P\rangle$, where S is a non-empty set and P is a binary relation on S, i.e., a subset of S^2 such that for any elements a, b, and c of S, the following four conditions are satisfied:

L1. $\langle a,a\rangle \notin P$ (P is irreflexive).

L2. If $\langle a,b\rangle \in P$ and $\langle b,c\rangle \in P$, then $\langle a,c\rangle \in P$ (P is transitive).

L3. Either $\langle a,b\rangle \in P$ or $\langle b,a\rangle \in P$ or $a = b$ (P is connected).

L4. $\langle a,b\rangle \in P$ entails that $\langle b,a\rangle \notin P$ (P is asymmetric).

P is called the ordering relation and is usually symbolized by $<$ (or by $>$). One writes $a < b$ (or, alternatively, $a > b$) for $\langle a,b\rangle \in P$.

If the following conditions are substituted for L1–L4, $\langle S,P\rangle$ is said to be a *weak order*.

W1. $\langle a,a\rangle \in P$ (P is reflexive).

W2. If $\langle a,b\rangle \in P$ and $\langle b,c\rangle \in P$, then $\langle a,c\rangle \in P$ (P is transitive).

W3. Either $\langle a,b\rangle \in P$ or $\langle b,a\rangle \in P$ (P is strongly connected).

W4. Unless $a = b$, $\langle a,b\rangle \in P$ entails that $\langle b,a\rangle \notin P$ (P is antisymmetric).

The weak ordering relation P is usually symbolized by $\leq$ (or by $\geq$). We write $a \leq b$ (or, alternatively, $a \geq b$) for $\langle a,b\rangle \in P$. If P is required to be reflexive, transitive and antisymmetric, but neither connected nor strongly connected, $\langle S,P\rangle$ is a *partial order*. The three kinds of order here defined have the following, easily provable, property: If $\langle S,P\rangle$ is a linear, weak, or partial order, there is an order $\langle S,P'\rangle$, of the same kind as $\langle S,P\rangle$, such that for any elements a and b of S, $\langle a,b\rangle \in P'$ if and only if $\langle b,a\rangle \in P$. This property justifies the dual notations, $(<, >)$ and $(\leq, \geq)$.

[For example, if S is the set of instants discerned in a given time interval I and P stands for temporal precedence, $\langle S,P\rangle$ is a linear order. If S is the set of all events beginning during the interval I and P denotes the relation 'x began before y', $\langle S,P\rangle$ is a weak order. If S is any set, $\langle \mathcal{P}(S), \subset\rangle$ is a partial order.]

A *field* is an octuple $\langle F,F',e,f,g,e',f',g'\rangle$, subject to the following requirements:

F1. F and F' are non-empty sets.

F2. $\langle F, e, f, g\rangle$ is an abelian group (g is called 'addition'; one writes '$a + b$', rather than $g(a,b)$; the neutral element e is usually called 'zero' and denoted by 0).

F3. F' is the complement of $\{e\}$ in F.

F4. g' is a mapping of F^2 into F such that for any $a \in F$, $g'(a,e) = g'(e,a) = e$ (g' is called 'multiplication'; one writes '$a \times b$' or simply 'ab', rather than $g'(a,b)$).

F5. g' is associative (in the sense explained in G1).

F6. $\langle F', e', f', g''\rangle$ is an abelian group, where g'' denotes the restriction of g' to F' (e' is usually called 'one' or 'unity' and denoted by 1).

F7. g and g' obey the following law of distribution: For any elements a, b, and c of F, $g'(a,g(b,c)) = g(g'(a,b),g'(a,c))$.

A field $\mathbf{F} = \langle F,F',e,f,g,e',f',g'\rangle$ is said to be *ordered* if it is divided into two mutually exclusive subsets, the *negative* and the *non-negative* elements of $\mathbf{F}$, in such a way that the neutral elements e and e' (zero and unity) are both non-negative, and that $g'(a,a)$—i.e., $a \times a$—is non-negative for any $a \in F$. It will be readily seen that if $\mathbf{F}$ is an ordered field in the sense just defined, its primary base set F is linearly ordered by the following condition:

For any $a,b \in F$, $a < b$ if and only if $a - b$ is negative (where $a - b$ stands for $a + f(b)$, i.e., $g(a,f(b))$).

Let $\mathbf{F} = \langle F, F', e, f, g, e', f', g' \rangle$ be an ordered field. If $a \in F$, I write $-a$ for $f(a)$. Put $|a| = -a$ if a is negative, and $|a| = a$ if a is non-negative. Now consider (infinite) sequences in **F**, i.e., mappings of the full set of natural numbers into the primary base set F. If α_k denotes the value assigned by such a mapping to an arbitrary natural number k, we typically denote the sequence in question by $\alpha_1, \alpha_2, \ldots$ or, for greater brevity, by (α_k). The sequence (α_k) is said to be a *null* sequence if for every ε in F there is a natural number N such that for any natural number $n > N$, $|\alpha_n| < \varepsilon$. Take any given $\alpha \in F$. The sequence (α_k) is said to *converge to the limit* α, if the sequence $(\alpha_k - \alpha)$ is a null sequence. (α_k) is said to be a Cauchy sequence if for every ε in F there is a natural number M such that for any natural numbers $m,n > M$, $|\alpha_m - \alpha_n| < \varepsilon$. The ordered field **F** is said to be *complete* if every Cauchy sequence in **F** converges to a limit in **F**.

If N is any positive integer, and $a \in F$, we write Na for $a + a + \ldots + a$, with a repeated N times (where '$a + a$' stands for $g(a,a)$). The ordered field **F** is said to be *Archimedean* if for any two non-negative elements $a,b \in F$, there always is a positive integer N such that $a < Nb$.

[There is an endless variety of fields, but two are of paramount importance in physics: the field **R** of real numbers and the field **C** of complex numbers.

To introduce **R**, one usually takes the field **Q** of rational numbers as the starting point. Let F denote the set of rationals (all negative and non-negative, proper and improper fractions). Let e be the rational number 0, and e' the rational number 1 (or 1/1). Let g and g' stand, respectively, for addition and multiplication of rationals, f for multiplication by -1, and f' for the operation of taking the reciprocal value. The reader should verify that, in this interpretation, the octuple $\langle F, F', e, f, g, e', f', g' \rangle = \mathbf{Q}$ is an ordered (in effect, Archimedean) field. Consider Cauchy sequences in **Q**. Many of them converge to a rational limit, but many more do not. For example, the sequence $((1 + 1/k)^k)$ does not converge to a rational limit. Thus **Q** is not a complete field.

The field **R** is defined from **Q** with the deliberate aim of filling the gaps that spoil the latter's completeness. Let us say that two sequences of rationals, (α_k) and (β_k), are equivalent if the sequence $(\alpha_k - \beta_k)$ is a null sequence. Let $[\alpha_k]$ stand for the equivalence class which contains the sequence (α_k). Reinterpret F to be the set of all equivalence classes of Cauchy sequences of rationals. We now define addition and multiplication on F by the rules $[\alpha_k] + [\beta_k] = [\alpha_k + \beta_k]$ and $[\alpha_k] \times [\beta_k] = [\alpha_k \times \beta_k]$. (Note that the symbols '+' and '×' inside the brackets represent addition and multiplication of rationals; the same symbols outside the brackets stand for the newly defined addition and multiplication of equivalence classes of Cauchy sequences of rationals. Of course, the definitions only make sense because it can be shown that if the sequence (α_k) is equivalent to (α'_k) and (β_k) is equivalent to (β_k'), then the sequence $(\alpha_k + \beta_k)$ is equivalent to $(\alpha_k' + \beta_k')$, and the sequence $(\alpha_k \times \beta_k)$ is equivalent to $(\alpha_k{}' \times \beta_k{}')$.) Let e be the equivalence class of the sequence of rationals $(0, 0, \ldots)$, and let e' be the equivalence class of the sequence of rationals $(1, 1, \ldots)$. The reader should satisfy himself that the octuple $\langle F,F',e,f,g,e',f',g'\rangle$, as now interpreted, is again an Archimedean field. We call it the field of real numbers and designate it by **R**. It can be shown—though this is not the place to do it—that every Cauchy sequence of real numbers converges to a limit in **R**. As desired, **R** is a complete field. Indeed, any complete Archimedean field is isomorphic (in the exact sense explained in Section 2.8.5) with the field **R** just defined. Since isomorphic structures are mathematically indistinguishable, the field **R** of real numbers may be straightforwardly characterized as the one and only complete Archimedean field. Thus, the most pervasive structure of mathematical physics can be introduced directly, without subordinating it to the rationals in the clever, yet undeniably farfetched way we have followed. It is true of course that the equivalence classes of Cauchy sequences of rationals form a complete Archimedean field, as described above, and are therefore an instance of **R**, which is isomorphic with and thus provides an optimal representation for the complete Archimedean fields of nature.

The field **C** of complex numbers may be defined by taking as the base set F the set of all ordered pairs of reals, and defining addition and multiplication by the rules: $\langle\alpha,\beta\rangle + \langle\gamma,\delta\rangle = \langle\alpha + \gamma,\beta + \delta\rangle$, and $\langle\alpha,\beta\rangle + \langle\gamma,\delta\rangle = \langle\alpha\gamma - \beta\delta, \alpha\delta + \beta\gamma\rangle$.]

A *vector space over the field* **F**, or an **F**-*vector space,* is a 13-tuple $\langle V,F,F',\mathbf{e},\mathbf{f},\mathbf{g},e,f,g,e',f',g',h\rangle$ satisfying the following conditions:

V1. V and F are non-empty sets, whose elements are called vectors and scalars, respectively.

V2. $\langle V,\mathbf{e},\mathbf{f},\mathbf{g}\rangle$ is an abelian group (**g** is usually called 'vector addition'; the neutral element **e** is called 'the zero vector' and denoted by **0**).

V3. $\langle F,F',e,f,g,e',f',g'\rangle$ is a field.

V4. h is a mapping of $F\times V$ into V (called 'multiplication by a scalar').

V5. For any scalar α and any vectors **v** and **w**, $h(\alpha,\mathbf{g}(\mathbf{v},\mathbf{w})) = \mathbf{g}(h(\alpha,\mathbf{v}),h(\alpha,\mathbf{w}))$.

V6. For any scalars α and β, and any vector **v**, $h(g(\alpha,\beta),\mathbf{v}) = \mathbf{g}(h(\alpha,\mathbf{v}),h(\beta,\mathbf{v}))$.

V7. For any scalars α and β, and any vector **v**, $h(\alpha,h(\beta,\mathbf{v})) = h(g''(\alpha,\beta),\mathbf{v})$.

V8. For any vector **v**, $h(e',\mathbf{v}) = \mathbf{v}$.

For the sake of perspicuity one usually writes $\mathbf{v}+\mathbf{w}$ for $\mathbf{g}(\mathbf{v},\mathbf{w})$, and $\alpha\mathbf{v}$ for $h(\alpha,\mathbf{v})$. One also writes $\mathbf{v} - \mathbf{w}$ for $\mathbf{v} + \mathbf{f}(\mathbf{w})$. Vector spaces are readily enriched with further structure. Consider the 14-tuple $\mathcal{V} = \langle V,\mathbf{R},\mathbf{R}\backslash\{0\},\mathbf{e},\mathbf{f},\mathbf{g},e,f,g,e,f',g',h,\varphi\rangle$, where $\langle V,\mathbf{R},\mathbf{R}\backslash\{0\},\mathbf{e},\mathbf{f},\mathbf{g},e,f,g,e',f',g',h\rangle$ is a vector space over $\mathbf{R}$ or real vector space and φ is a mapping of $V\times V$ into $\mathbf{R}$, such that (in the notation just described), for any $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, $\alpha, \beta \in \mathbf{R}$,

(i) $\varphi(\mathbf{u},\mathbf{v}) = \varphi(\mathbf{v},\mathbf{u})$,

(ii) $\varphi(\alpha\mathbf{u} + \beta\mathbf{v},\mathbf{w}) = \alpha\varphi(\mathbf{u},\mathbf{w}) + \beta\varphi(\mathbf{v},\mathbf{w})$,

(iii) $\varphi(\mathbf{u},\mathbf{u}) \geq 0$, and

(iv) $\varphi(\mathbf{u},\mathbf{u}) = 0$ if and only if $\mathbf{u} = \mathbf{0}$.

The mapping φ is called an *inner product* and $\mathcal{V}$ is said to be an *inner product space*. Instead of $\varphi(\mathbf{u},\mathbf{v})$ one writes $\langle\mathbf{u},\mathbf{v}\rangle$, or $\langle\mathbf{u}|\mathbf{v}\rangle$, or $\mathbf{u} \bullet \mathbf{v}$. The foregoing characterization of inner product spaces can be extended to complex vector spaces (i.e., vector spaces over the field $\mathbf{C}$) by substituting $\mathbf{C}$ for $\mathbf{R}$ wherever the latter symbol occurs in it, and replacing (i) with

(i*) $\varphi(\mathbf{u},\mathbf{v}) = (\varphi(\mathbf{v},\mathbf{u}))^*$ (where α^* stands for the complex conjugate of $\alpha \in \mathbf{C}$).

The following ideas will be of use in Section 5.3. Let $\mathcal{V}$ be a real or complex inner product space, as described above. The mapping of V into $\mathbf{R}$ by $\mathbf{u} \mapsto |\varphi(\mathbf{u},\mathbf{u})|^{1/2}$ is called the *norm* on $\mathcal{V}$. One writes $\|\mathbf{u}\|$ instead of $|(\varphi(\mathbf{u},\mathbf{u})|^{1/2}$. $\|\mathbf{u}\|$ is also called the *length* of the vector $\mathbf{u}$. Consider an infinite sequence of vectors, $\mathbf{u}_1, \mathbf{u}_2, \ldots \in V$. We say that $\mathbf{u}_1, \mathbf{u}_2, \ldots$ is a Cauchy sequence if for every positive real number ε there is a positive integer N such that for all integers $m,n > N$ we have that $\|\mathbf{u}_m - \mathbf{u}_n\| < \varepsilon$. The inner product space $\mathcal{V}$ is said to be a (real or complex) *Banach space* if every Cauchy sequence $\mathbf{u}_1, \mathbf{u}_2, \ldots \in V$ converges to a vector in V, i.e., if for each such sequence there is a fixed vector $\mathbf{u} \in V$ which is such that for every positive real number ε there is a positive integer N such that $\|\mathbf{u} - \mathbf{u}_n\| < \varepsilon$ for every $n > N$.

The nth Cartesian product $\mathbf{R}^n$ of the real number field with itself can be seen in a natural way as a real vector space, with the following obvious definitions of (1) vector addition and (2) multiplication by a scalar:

(1) $\langle u_1, \ldots, u_n\rangle + \langle v_1, \ldots, v_n\rangle = \langle u_1 + v_1, \ldots, u_n + v_n\rangle$;

(2) $\alpha\langle u_1, \ldots, u_n\rangle = \langle \alpha u_1, \ldots, \alpha u_n\rangle$.

The zero vector is plainly the n-tuple $\langle 0, \ldots, 0\rangle$. The standard inner product on $\mathbf{R}^n$ is given by

(3) $\langle u_1, \ldots, u_n\rangle \bullet \langle v_1, \ldots, v_n\rangle = u_1v_1 + \ldots + u_nv_n$.

This clearly yields the so-called Euclidian (or Pythagorean) norm:

(4) $\|\langle u_1, \ldots, u_n \rangle\| = u_1{}^2 + \ldots + u_n{}^2$.

The reader should verify that, with this structure, $\mathbf{R}^n$ is indeed a Banach space.

2.8.5 Isomorphism

Let us now define the much abused term 'isomorphism'. Consider two structures $\langle A_1, \ldots, A_r, a_1, \ldots, a_m \rangle$ and $\langle B_1, \ldots, B_s, b_1, \ldots, b_n \rangle$, with base sets $A_1, \ldots, A_r$ and $B_1, \ldots, B_s$, respectively, and a mapping $\mathbf{h}$ from $A_1 \cup A_2 \cup \ldots \cup A_r$ to $B_1 \cup B_2 \cup \ldots \cup B_s$. $\mathbf{h}$ is an *isomorphism* between the said structures if and only if (i) $\mathbf{h}$ is bijective; (ii) $r = s$ and $m = n$; (iii) for some permutation π of $\{1, \ldots, r\}$, the restriction of $\mathbf{h}$ to A_k is a bijection onto $B_{\pi k}$ $(1 \leq k \leq r)$; and (iv) the homonymous mappings induced by $\mathbf{h}$ in the echelon sets over A and by the inverse mapping $\mathbf{h}^{-1}$ in the echelon sets over B are such that for some permutation σ of $\{1, \ldots, m\}$, with inverse permutation ρ, $\mathbf{h}(a_i) = b_{\sigma i}$ and $\mathbf{h}^{-1}(b_i) = a_{\rho i}$ $(1 \leq i \leq m)$. Note that conditions (ii) and (iv) jointly imply that $\langle A_1, \ldots, A_r, a_1, \ldots, a_m \rangle$ and $\langle B_{\pi 1}, \ldots, B_{\pi s}, b_{\sigma 1}, \ldots, b_{\sigma n} \rangle$ are, in effect, two similar (r,m)–lists of structural components (see Section 2.8.4). Two structures are *isomorphic* if there exists an isomorphism between them.

[In the first example of of Section 2.8.4 I represented the boy-girl situation of Subsection (b) by the structure $\langle B, G, <_B, <_G, f \rangle$. Now, let Algernon, Angus and Marmaduke be three gluttons, such that Algernon dines on Petunia the cow and Jason the elephant, Angus devours Jonathan the rhinoceros and Edsel the whale, while Marmaduke sleeps. Let G' be the set of gluttons, B' the set of beasts, and f' the mapping that assigns to each animal the glutton who eats it. Let $<_{B'}$ be a linear order on B' in agreement with the alphabetical order of the animal species involved, and $<_{G'}$ a linear order on G' in agreement with the alphabetical order of individual names. Then, I claim, $\langle G', B', f', <_{B'}, <_{G'} \rangle$ is isomorphic with $\langle B, G, <_B, <_G, f \rangle$. To prove it, let $\mathbf{h}$ map $B' \cup G'$ onto $B \cup G$ by sending each element x of the domain to the element of the codomain which shares with x the first two letters of its name. If we choose $(1,2) \rightarrow (2,1)$ to be the permutation π, and $(1,2,3) \rightarrow (3,1,2)$ to be the permutation σ, it will be clear that $\mathbf{h}$ satisfies the conditions prescribed above for an isomorphism.]

2.8.6 Alternative typifications

Consider now two species of structure, $\Sigma_1 = \langle S_1, \ldots, S_m, A_1, \ldots, A_h \rangle$ and $\Sigma_2 = \langle S_1, \ldots, S_m, B_1, \ldots, B_k \rangle$, which share the same list of base sets $\mathfrak{S} = \langle S_1, \ldots, S_m \rangle$ and are governed by conditions $\mathcal{H}$ and $\mathcal{K}$, respectively. Let us denote by $\mathcal{H}'$ and $\mathcal{K}'$ the sentences formed by priming the names of the structural components $A_1, \ldots, A_h, B_1, \ldots, B_k$ in $\mathcal{H}$ and $\mathcal{K}$, respectively. Suppose that from apposite echelon sets over $\mathfrak{S}$ we can select elements $B'_1, \ldots, B'_k$ by virtue of $\mathcal{H}$, and elements $A'_1, \ldots, A'_h$ by virtue of $\mathcal{K}$, so that $\langle S_1, \ldots, S_m, B'_1, \ldots, B'_k \rangle$ satisfies conditions $\mathcal{K}'$ and $\langle S_1, \ldots, S_m, A'_1, \ldots, A'_h \rangle$ satisfies conditions $\mathcal{H}'$. Plainly, any particular instance of either type of structure also provides an instance of the other. We therefore regard Σ_1 and Σ_2 as equivalent species of structure, or, better still, as alternative typifications of what is at bottom the same kind of mathematical object, viz., the same species of structure.

I shall illustrate this idea by giving three alternative typifications of the important species of structure known as a topological space. By reflecting on them one can better understand why the identity of a species of structure cannot be made to rest on a particular formal description. The second typification is especially elegant as it characterizes a topological space by means of a simple "operation" on the power set of its base set. A brief and very lucid motivation of the other two will be found in the splendid book by Saunders Mac Lane (1986, pp. 30–33).

A *topological space* is a pair $\langle S, \mathbf{T} \rangle$, where S is a set and $\mathbf{T}$ is a subset of $\mathcal{P}(S)$—and thus a distinguished element of $\mathcal{P}^2(S)$—which satisfies the following three conditions:

T_O1. S belongs to $\mathbf{T}$.

T_O2. The intersection of any two elements of $\mathbf{T}$ belongs to $\mathbf{T}$. ($A, B \in \mathbf{T} \Rightarrow A \cap B \in \mathbf{T}$.)

T_O3. The union of any subset of $\mathbf{T}$ belongs to $\mathbf{T}$. ($\mathbf{H} \subset \mathbf{T} \Rightarrow \bigcup \mathbf{H} \in \mathbf{T}$.)

$\mathbf{T}$ is called a *topology* on S. An element of $\mathbf{T}$ is said to be an *open set*. If x is an element of S, any subset of S that includes an open set which contains x is said to be a *neighborhood* of x. If $A \subset S$ and A is open, the complement of A in S is said to be *closed*. The mapping f: $\mathcal{P}(S) \to \mathcal{P}(S)$ that sends each subset A of S to the intersection of all closed sets in which A is included is the *closure* of $\langle S, \mathbf{T} \rangle$. Note that T_O3 entails that $\varnothing$ is open, and so, by T_O1, S and $\varnothing$ are both open and closed. $\mathbf{T}$ is the *trivial* topology on S if $\mathbf{T} = \{S, \varnothing\}$. $\mathbf{T}$ is the *discrete* topology on S if $\mathbf{T} = \mathcal{P}(S)$.

Alternatively, a *topological space* is a pair $\langle S, f \rangle$, where S is a set and f is a mapping of $\mathcal{P}(S)$ into itself such that for any subsets A and B of S, the following four conditions are met:

T_K1. $f(A \cup B) = f(A) \cup f(B)$.

T_K2. $A \subset f(A)$.

T_K3. $f(f(A)) = f(A)$.

T_K4. $f(\emptyset) = \emptyset$.

f is said to be the *closure* of S. Any set lying on its range is said to be *closed.* If A is closed in S, the complement of A in S is said to be *open.* The collection of the open sets of S is the topology of $\langle S,f\rangle$. Neighborhoods are defined as above. It can be easily proved that if f denotes the closure of a topological space $\langle S,\mathbf{T}\rangle$, characterized by conditions T_O, f meets the conditions T_K, and that if T denotes the topology of a topological space $\langle S,f\rangle$, characterized by conditions T_K, T meets the conditions T_O.

But a topological space may be characterized in still another way. Let S be a set and ϕ a mapping of S into $\mathcal{P}^2(S)$ that assigns to each element x of S a collection $\phi(x)$ of subsets of S meeting the following conditions:

T_N1. $S \in \phi(x)$.

T_N2. $x \in U$, for every $U \in \phi(x)$.

T_N3. If $U \in \phi(x)$ and $U \subset V$, $V \in \phi(x)$.

T_N4. If $U \in \phi(x)$ and $V \in \phi(x)$, $U \cap V \in \phi(x)$.

T_N5. If $U \in \phi(x)$, there is a set V such that $x \in V \subset U$ and $V \in \phi(y)$ for every $y \in V$.

Then, $\langle S,\phi\rangle$ is a *topological space.* If $x \in S$, every element of $\phi(x)$ is a *neighborhood* of x. A subset of S is said to be *open* if it is a neighborhood of every element it contains. (Note that this condition is satisfied by $\emptyset$, for it contains no elements.) A subset of S is said to be *closed* if its complement in S is open. The collection of all open sets is the *topology* of $\langle S,\phi\rangle$. The mapping that sends each subset of S to the intersection of all closed sets that include it is the *closure* of $\langle S,\phi\rangle$. It can then be shown that if $\mathbf{T}$ is the topology of $\langle S,\phi\rangle$, $\mathbf{T}$ meets the conditions T_O; and that if f is the closure of $\langle S,\phi\rangle$, f meets the conditions T_K. On the other hand, if we choose to understand 'neighborhood' in the sense defined above after the statement of conditions T_O, and $\phi(x)$ denotes the collection of the neighborhoods of an arbitrary $x \in S$, it can be shown that $\phi(x)$ meets the conditions T_N.

[Let me illustrate the above ideas in the light of the standard topology of the Euclidian plane **E**. A set of points of **E** whose distance from a point P is less than a given real number is said to form an open disk at P. A set of points whose distance from P is not greater than a given real number is said to form a closed disk at P. For greater precision, writing $|X - Y|$ for the distance from point X to point Y, I define the open disk (P,ρ), with center P and radius ρ, as the point set $\{X \mid |X - P| < \rho\}$, and the closed disk $[P,\rho]$, with center P and radius ρ, as the point set $\{X \mid |X - P| \leq \rho\}$ ($P \in E$, $0 \leq \rho \in \mathbf{R}$). A point set is open in the standard topology of the Euclidian plane if (i) it is an open disk, or (ii) it is the intersection of two open sets, or (iii) it is the union of a family of open sets. A point set is closed in that topology if (i) it is a closed disk, or (ii) it is the union of two closed sets, or (iii) it is the intersection of a family of closed sets. Let $P \in \mathbf{E}$ and let ζ be a subset of **E** that includes some open disk at P; then, ζ is a neighborhood of P. A closure mapping is defined on the Euclidian plane by assigning to each point set the intersection of all closed sets that include it. It can also be characterized by defining the mapping $(P,\rho) \mapsto [P,\rho]$ on the set of open disks and extending this mapping to all subsets of **E** in such a way that conditions T_K1–T_K4 are all met. (It must of course be shown that there is a unique such extension.)]

2.8.7 Axiomatic set theory

Bourbaki's *Éléments de Mathématique* has shown that it is possible to conceive the familiar areas of mathematical inquiry as the study of different species of structure (and their subspecies). Of course, Bourbaki did not follow the naïve approach to sets adopted above. As I noted in Section 2.8.1, if we nonchalantly regard any expression of the form '$\{x \mid S(x)\}$' as the description of a set, we run into contradictions. To avoid them, Russell (1908) sought to regiment the language of science, drastically and, some would say, intemperately restricting the admissible expressions. But working mathematicians have generally preferred to deal with this problem after the manner of Zermelo (1908b), who proposed to characterize sets by an axiom system free from contradiction and not to countenance any set descriptions not warranted by this axiom system. Zermelo's axioms, subsequently perfected by Fraenkel (1922, etc.) and known thereafter as the ZF axioms, assert that certain sets exist *if* certain other sets are given (two of these conditional statements were quoted as (i) and (ii) in Section 2.8.3), and postulate that there is at least one infinite set. It is demonstrably impossible to prove the consistency of the ZF axioms without invoking premises at least as strong. But hitherto no contradictions have cropped up in axiomatic set theory. However, the ZF axioms do not by themselves suffice to characterize the mathematical concept of a set. Zermelo himself (1904, 1908a) had used one additional assumption—the so-called Axiom of Choice[77]—for proving that every set can be well-ordered.[78] Gödel (1938, 1940) proved that the ZFC axioms (Zermelo-Fraenkel axioms plus the Axiom of Choice) are consistent *if* the ZF axioms are consistent. But then Paul Cohen (1963/64) proved that the ZF¬C axioms (Zermelo-Fraenkel axioms plus the negation of the Axiom of Choice) are consistent too if the ZF axioms are consistent. Do we mean by 'sets' the objects axiomatically characterized by the ZFC system or those characterized by the ZF¬C system? Most mathematicians will not hesitate to choose the former, for far too many strong and beautiful mathematical results cannot be proved in the latter. R. M. Solovay (1970) has shown that the best known of them follow also from a weakened and in some ways more convenient version of the Axiom of Choice. However, Solovay's proposal raises the specter of some very large sets, whose existence not everyone will concede. (Cf. Section 4.2, note 7.)

A similar question arises in connection with Cantor's Continuum Hypothesis and the Generalized Continuum Hypothesis, which I shall here denote by H^c and H^g, respectively. The import of H^c and H^g can be simply explained as follows. Let us say that a set A is *strictly smaller* than a set B—and B *strictly larger* than A—if there is an injective mapping of A into B, but it is impossible to map B injectively into A. Cantor proved that every set is strictly smaller than its power set. Let us say that the power set $\mathcal{P}(A)$ of a set A is *next to* A if it does not include any subset that is both strictly smaller than $\mathcal{P}(A)$ and strictly larger than A. It is plain that if a set A has only one element, $\mathcal{P}(A)$ is next to it. However, if A is finite, but has more than one element, $\mathcal{P}(A)$ is not next to A. For example, if A has 3 elements, $\mathcal{P}(A)$ has 8 elements and includes subsets with 4, 5, 6, and 7 elements. Cantor, however, expected to prove that, if $\mathbf{N}$ denotes the set of the natural numbers, $\mathcal{P}(\mathbf{N})$ is next to $\mathbf{N}$. This is Cantor's Continuum Hypothesis H^c. The Generalized Continuum Hypothesis H^g asserts that if A is any infinite set, $\mathcal{P}(A)$ is next to A. Gödel (1938, 1940) proved that ZFH^g is consistent if ZF is consistent. But again Cohen (1963/64) proved that ZF¬H^c—and hence ZF¬H^g—is consistent, if ZF is consistent. When we say that a group or a topological space is a *set* endowed with such-and-such structural components, do we mean that it satisfies H^c or that it satisfies ¬H^c?

2.8.8 Categories

The concept of a mathematical category was first introduced by Eilenberg and Mac Lane (1945). Category theory provides a global approach to mathematics that is more flexible, closer to real mathematical work and possibly more fruitful than the exclusively set-theoretic perspective adopted by Bourbaki. While the Bourbaki group defines a species of mathematical structure by the distinctive internal properties and relations of a typical representative (e.g., a topology on an arbitrary set), the new school of thought envisions any specific field of mathematics (e.g., topology) as a system of objects (e.g., the topological spaces) characterized by the peculiar nature of the network of structurally significant mappings from one to the other (e.g., continuous functions, homeomorphisms). I will not go further into this matter.[79] Joseph Sneed and some of his followers have been using category-theoretic concepts in some of their more recent work on physical theories, but the book by Balzer, Moulines, and Sneed (1987) on which I shall base my discussion of Sneed's views in Chapter 3 describes all mathematical concepts in a strictly Bourbakian manner.

5. Necessity

We saw in Chapter 1 that impersonal observation, as currently practised in the natural sciences, can supply information about the states and evolution of observed objects only if—or insofar as—the latter constitute necessary conditions of the recorded receiver states. It is therefore apparent that the natural sciences can thrive only in a setting strung together by necessary connections. I noted at the end of Section 2.1 that inferential explanation—at least in its strong deductive-nomological form—provides just that, as it imbeds the explained phenomena in a domain of objects linked by bonds of necessity. Thus the construction of such bonds does not merely cater to the cravings of Reason but is also required for collecting the empirical data that go into the fabrication of scientific facts. Our subsequent explorations—above all in Chapter 3—should have thrown some light on the intellectual means deployed for that purpose by physics. But nothing has yet been said about the kind of necessity that physics discloses in natural phenomena. The subject is notoriously thorny—so much so, that some philosophers eschew physical necessity altogether. I should not say that its condition has improved with the recent attempts to explicate necessity and the twin concept of possibility in terms of a plurality of worlds.[1] I shall briefly refer to such attempts by way of introduction to the informal review of ordinary forms of possibility and necessity that will occupy us in Section 5.1. Then I shall discuss the Greek discovery of necessity in geometry in Section 5.2, the natural laws in the guise of differential equations that are the backbone of physics in Section 5.3, and the uneasy coexistence between the type of necessary connection implied by such laws and the ordinary idea of cause and effect in Section 5.4.

5.1 Forms of necessity

The concepts of possibility and necessity, as applied to sentences or statements or to the truths and falsehoods ("propositions") conveyed by them, play the role of syntactic/semantic operators, which map sentences to sentences and propositions to propositions, just as negation maps a sentence to its denial or conjunction maps any pair of sentences to their joint assertion. In this role they are known as the modal operators—or the classical modal operators—and their close mutual relationship can be lucidly displayed with the symbols of modern logic. I choose the symbols $\neg$ (read: 'it is not the case that'), $\Box$ ('it is necessary that'), and $\Diamond$ ('it is possible that') to represent negation, necessity, and possibility, respectively, and I let the letter **p** stand for any sentence in the indicative mood (the sort of sentence you can use to make a statement, capable of truth or falsity). All philosophers agree that $\Box$**p** if and only if $\neg\Diamond\neg$**p** and that $\Diamond$**p** if and only if $\neg\Box\neg$**p.** Following the fashion of formal semantics for the (nonmodal) predicate calculus, the polycosmist philosophers link the truth-conditions of $\Diamond$**p** and $\Box$**p** to the truth-values of **p** alone. Now, we may readily grant that it is possible that **p** if '**p**' is true[2] and that it is not necessary that **p** if '**p**' is false; but obviously being told that it is not the case that **p** will not teach us anything about the truth-value of '$\Diamond$**p**', nor will knowledge that **p** is the case give us an indication as to the truth-value of '$\Box$**p**'. Thus, for someone trained to equate the world with the aggregate of everything that is the case, the truth-conditions for sentences governed by the classical modal operators must be found outside it—whence the need for many worlds so that modal talk will make sense in ours. In polycosmist semantics, '$\Diamond$**p**' is true if and only if there is a world in which it is the case that **p**, and '$\Box$**p**' is true if and only if **p** is the case in every world.[3]

To the uncommitted observer this will seem a classic instance of philosophy getting mired in its own trash. Is the world no more than the aggregate of facts? Is understanding a sentence tantamount to knowing its truth-conditions?[4] Statements of possibility and necessity have been made and understood in every language—presumably since the dawn of man—without even the remotest hint of a plurality of worlds. Of course, the clever polycosmist will disclaim any intention to teach people how to use such ordinary modal operators as 'I can' or 'I have to', or to tell them what they are trying to say by means of them. His semantic rules define the modal operators contextually, i.e., they provide for every statement that contains a given modal operator another statement that does not contain it, yet allegedly always shares the truth-value of the former statement. Such definitions "in use" have a practical purpose when the definiens is more readily testable than the unaided definiendum. But the statements in question stand to each other precisely in the opposite relation. Since the definiens speaks of other worlds causally unconnected with our own, its truth-value must be taken on faith. However, it can be instantly computed from the known truth-value of the definiendum on the strength of the definition itself. For example, I know—of course, fallibly—that by pressing certain keys I can now type the

Greek letter sigma into the electronic manuscript I am writing. Hence, if I were to believe the polycosmist, I should rest assured that there is in another world a creature working on a device indistinguishable from a Macintosh computer whose life up to a certain time *t* is equal—as far as it and I and our respective friends can tell—to what has hitherto been my life, but which at time *t* does what I shall not do now, viz., to simultaneously press the keys command-shift-Q followed by S. When trying to figure out what I can or cannot do in a particular situation, so as to act accordingly, this sort of assurance leaves me cold. Polycosmist beliefs—or should we say conventions?—may favor the advancement of some theoretical projects,[5] but otherwise they are of no consequence. After all, what happens in *some* world will only matter to us if that world *can* be this one.

One of the polycosmist's chief worries is that rejection of his definitions will force us to take 'possible' or 'necessary' as primitive concepts. I say, what of it?[6] Although modal statements are often uncertain, this does not imply that their import is unclear. I find that our entire lives turn on our awareness of what is and what is not possible, no less than on our awareness of what is or has been actual. Indeed, much of our attention to the details of reality is prompted by our abiding quest for avenues—and barriers—to possibilities. A sense for the possible—and the impossible—is a major ingredient of the conscious performance of even the tiniest actions, e.g., bending the index finger or opening the mouth (faster? slower? more? less?—I sense it is up to me, within limits that I also sense). In fact, when no possibilities are sensed, one may not be said to be aware of *acting*. I dare say that this minute sense of what I can and cannot do right here and now, a sense that is sometimes hesitating but often cocksure, that is mostly hushful but never dormant, lies at the heart of my understanding of possibility and necessity. Plainly, it stands on its own no less than my sense of reality. Or rather, it is built into the latter. For if what I have just said is true, I sense my very being as a hub of possibilities. They mainly concern my worldly wheelings and dealings and are enabled and constrained by possibilities and necessities residing in the world.[7] Many of these I infer or surmise or accept on trust. Others, as the saying goes, I have never even dreamt of. But there are possibilities and necessities that I perceive in things no less immediately than their so-called sense qualities. For instance, as I press and roll a lump of wax between the palms of my hands I feel that it can take a different shape. However, if the wax were cold, I would feel no less clearly that it cannot, at least not by pressing it and rolling it with my hands. And, if I hugged it for a while, I would perceive, as it warms up, how it gradually acquires that possibility. No doubt it will be objected that what I perceive if the wax is not cold is that it *actually* changes its shape, not that it *can* assume another one. To this I reply that if, as I presume, we do *sense* change, and not just *infer* its occurrence from the sequential perception of different states of rest, we perceive in each change the possibility which it enacts much as we perceive our own possibilities in the course of action.[8]

The possibilities disclosed in such elementary experiences, or constructed from them in thought, may be termed *powers*—in good agreement with ordinary usage.[9] The corresponding impossibilities are often lived as powerlessness, for instance, when one watches the milk spilling from an overturned glass or tries in vain to stop a skidding automobile from crashing against another one. Our primary paradigm of necessity, brutal or coercive necessity, bemoaned by the poets and familiar to all, is rooted in such experiences. Philosophical attempts to reduce it to the purportedly more primitive notion of reality or existence seem to me extravagant, inasmuch as necessity thus perceived often constitutes an important criterion or indication of real existence. Is not our inability to abolish a fact a main ingredient of our awareness of its being factual? It is also worth noting how the effectiveness of power rests on the workings of necessity. I *can* unscrew a tight screw because the latter *must* yield to the torque transmitted by the screwdriver.

The common notions of possibility and necessity, as they show up in the ordinary usage of expressions that convey them, display features often ignored in philosophical discussions. Of course, this may be taken to mean simply that the philosophical concepts that are being discussed differ from their vulgar homonyms. I believe we ought to be aware of such differences and try to understand their motivation. Do they respond solely to a desire to trim off inconsistencies found in the common notions? Or did they arise in the course of extending these notions to a context for which they were not originally meant? I shall now review some peculiarities of the ordinary usage of modal expressions and speculate a little about the paths that may have led from it to the usage of philosophers.

Let me note, first of all, that in ordinary conversation we would never acknowledge a power or a possibility where there are no alternatives. You can easily fall while jogging on an uneven road, but not after you have already stumbled and are lying on the ground (unless, indeed, the ground itself lies at the top of a precipice from which you can go on falling). On the other hand, in philosophy whatever is known to be necessary is also said to be possible. This philosophical usage can trace its pedigree to Aristotle. However, as Jaakko Hintikka (1960) has pointed out, Aristotle often speaks on these matters in perfect agreement with our ordinary usage. For him, 'to be possible' involves a twofold possibility: if a thing may be, it may also not be.[10] Indeed, in a passage to which he subsequently refers as a definition (διορισμός) of possibility, he explicitly separates the necessary from the possible.

> λέγω δ' ἐνδέχεσθαι καὶ τὸ ἐνδεχόμενον, οὗ μὴ ὄντος ἀναγκαίου τεθέντος δ' ὑπάρχειν, οὐδὲν ἔσται διὰ τοῦτ' ἀδύνατον.
>
> I use the terms 'possibly' and 'the possible' of that which is not necessary but, being assumed, results in nothing impossible.
>
> (*An. Pr.* I, 13, $32^{a}18$–20)

Aristotle admits that the necessary is sometimes said to be possible *by homonymy.*[11] But, strictly speaking, possibility and necessity are to be distinguished as two mutually exclusive and jointly exhaustive modes of being:

> ἔσται ἄρα τὸ ἐνδεχόμενον οὐκ ἀναγκαῖον καὶ τὸ μὴ ἀναγκαῖον ἐνδεχόμενον.
>
> That which is possible, then, will not be necessary; and that which is not necessary will be possible.
>
> (*An. Pr.* I, 13, 32a28f.)
>
> καὶ ἔστι δὴ ἀρχὴ ἴσως τὸ ἀναγκαῖον καὶ μὴ ἀναγκαῖον πάντων ἢ εἶναι ἢ μὴ εἶναι.
>
> And, indeed, the necessary and not necessary are perhaps the principle of everything's either being or not being.
>
> (*De Int.* 13, $23^{a}18$f.)

Now, if every possibility involves a plurality of options, there are no possibilities in what is already past. Aristotle did not hesitate to draw this conclusion:

> οὐδεμία γὰρ δύναμις τοῦ γεγονέναι ἐστίν, ἀλλὰ τοῦ εἶναι ἢ ἔσεσθαι.
>
> There is no power of having been, but only of being or going to be.
>
> (*De Caelo* I, 12, $283^{b}13$)

Whence, of course, "what has already happened is necessary" (ἔχει γὰρ τὸ γεγονὸς ἀνάγκην—*Rhet.* III, 17, $1418^{a}3$-5; see also *Eth. Nich.* VI, 2, $1139^{b}7$–11). In this view, the past and the future do not just differ chronologically—in that they respectively precede and follow a particular moment of time, which happens to be now—but are two essentially distinct realms of being, one of which is being gradually transmuted into the other all the time. The limit between them—τὸ νῦν, the now—is the locus of fading options, where protean possibilities are ever taking a definite and definitive—necessary—shape.

Most philosophers find such metaphysics unpalatable, but it is certainly close to common sense. However, in everyday talk one often speaks of possibilities concerning the past. I do not mean just past possibilities, i.e., options which were once open but are now closed—for instance, the possibility of preventing World War I through diplomatic channels in July 1914. When such possibilities were in effect, they concerned the future and do not therefore constitute a problem for the stated view. But we also say quite commonly that it *is* possible—now—that such-and-such *was* the case—in the past. We may thus mention the possibility that X.Y.Z. murdered U.V.W., or that the Solar System resulted from a close encounter between two stars. Should we therefore conclude that the above views on modality and time do not, after all, agree with common sense? Questions about the contents of common sense are notoriously slippery, for any attempt to answer them naturally requires more decision and precision than one may wisely ascribe to it. However, it may be instructive to note that (i) we would not say that *it is possible* that X.Y.Z. murdered U.V.W. if we knew for certain that he did so, and (ii) we would regard the said modal statement as refuted if we learn, for instance, that X.Y.Z. was in New Zealand when U.V.W. was stabbed in Beirut. It appears, therefore, that when we talk of present possibilities with regard to past events all we mean to imply is that there are alternative ways in which our incomplete knowledge of such events may eventually be completed, not that there are any options open as to how they actually took place (or as to how their "counterparts" happen in different worlds).

By lifting our attention from particulars—be they individual events, processes, situations, states, or things—to their respective types, we can derive from the ordinary timebound notion of possibility as power a time-independent concept somewhat more akin to the possibility of philosophers. A natural way of doing it would be the following. Let us say that a type K of objects is possible if and only if there was or will be a time at which some object of type K is possible in the ordinary, timebound sense—a time, that is, at which the power exists to produce or not to produce such an object. Then, K is impossible, and the complementary type K′ is necessary, if and only if there never is a power to produce an instance of K. The possibility and necessity we have thus defined are not the timeless modalities of philosophy, but their ascription, if true, would not change with time. Evidently, a necessary type is actualized at all times and is, in this sense, eternal. But, of course, it does not follow that everything eternal is therefore necessary. Curiously, however, Aristotle draws just this conclusion:

> τὸ γὰρ ἐξ ἀνάγκης καὶ ἀεὶ ἅμα· ὅ γὰρ εἶναι ἀνάγκη οὐχ οἷον τε μὴ εἶναι· ὥστ' εἰ ἔστιν ἐξ ἀνάγκης, ἀΐδιον ἐστι, καὶ εἰ ἀΐδιον, ἐξ ἀνάγκης.

> That which is *of necessity* is at the same time *always*. For what must necessarily be is incapable of not being. So that if something is of necessity, then it is eternal, and if eternal, then of necessity.
> (*De Gen. et. Corr.* II, 11, 337^b35–338^a2)

Now, if "capable of not being" (οἷον τε μὴ εἶναι) is understood as I have proposed, eternity does not entail necessity. A nuclear war is clearly possible even if in the end—viz., after mankind becomes extinct due to other causes—it turns out that there never was one. Aristotle's reasoning is valid, however, if a different concept of possibility is presupposed. The following Diodoran notion has been ascribed to him by Hintikka (1957, 1973) and others: A type K is possible if and only if there was or will be a time at which an instance of K is actual.[12] Suppose that K is, in this sense, not possible. Then, no instance of K is ever realized. Therefore, the complementary type K' is always actual. At the same time, since K is impossible, K' is necessary. In other words, a type K' is necessary if and only if it is actualized at all times. There are several passages in which Aristotle appears to endorse Diodoran modality. According to him, "it is evident that it cannot be true to say: 'this is possible, but it will not be'" (δυνατὸν μὲν τοδί, οὐκ ἔσται δέ—*Metaph.* IX, 4, 1047^b4–5). Also, "it is obviously impossible for that which is destructible not to be destroyed sometime" (ἀδύνατον φθαρτὸν ὄν μὴ φθαρῆναί ποτε—*De C.* I, 12, 283^a25). Other passages provide evidence to the contrary. The most remarkable among them concerns a cloak which it is possible to cut—or not to cut—but is never cut because it first wears out (*De Int.* 9, 19^a13ff.). Hintikka, however, does not think that this passage is at all relevant, because cutting the cloak is a particular event, and the concept under discussion applies to event types. Be that as it may, the issue is of no great consequence to us here, for the modalities of current philosophy are not Diodoran, and Diodoran modalities are blatantly at odds with ordinary usage. (Surely one would normally say of any given class of deployed nuclear missiles that it has posed a real danger, a possibility of violent death for millions of people, even if it grows obsolete and is decommissioned before ever being used.)

I surmise that to go from ordinary to philosophical modal discourse you do not have to climb from tokens to types, from particulars to universals, but that, as in other cases in which philosophers use common words in an odd way, you will find the missing link in medieval theology. The idea of God's power, or, rather, of God *as* power, is central to Judaic, Christian, and Islamic theology.[13] Such power is boundless, and therefore, in a sense, nothing should be impossible for Him. The theologian is required to explain exactly in what sense this is true. For surely God cannot destroy Himself or detract from His goodness. It will not do to say that God can do whatever is possible in the time-independent sense proposed above, i.e., whatever lies within anyone's or anything's capacity at any time and any place. God's omnipotence is more than just the sum total of all worldly powers.[14] On the other hand, we do not want a circular explanation that simply equates what is possible for God with what lies in His own power. To solve this conundrum, Aquinas resorts to the following concept of absolute possibility:

> Relinquit igitur quod Deus dicatur omnipotens, quia potest *omnia possibilia absolute,* quod est alter modus dicendi *possibile.* Dicitur autem aliquid possibile vel impossibile absolute, ex habitudine terminorum: possibile quidem, quia praedicatum non repugnat subiecto, ut Socratem sedere; impossibile vero absolute, quia praedicatum repugnat subiecto, ut hominem esse asinum.
>
> It remains to say that God is omnipotent because he can do everything that is absolutely possible. This is another sense of 'possible'. Something is said to be absolutely possible or impossible because of the disposition of terms; viz., possible, if the predicate is compatible with the subject, as in 'Socrates is seated'; but absolutely impossible, if the predicate is incompatible with the subject, as in 'the man is an ass'.
>
> (*Summa Theol.* 1, qu.25, a.3)

This is, of course, the concept of possibility that has prevailed in modern philosophy. According to Aquinas, it agrees with one of several meanings of 'possible' distinguished by Aristotle, viz., "the possible conceived without reference to a power" (δυνατά λεγόμενα οὐ κατὰ δύναμιν).[15] Next to the passage I have just quoted, and again, more clearly and at greater length, in the *Summa contra Gentiles,* 2.25, Aquinas shows why the word must be used in this particular acceptation to explain what is or is not possible for God. The gist of his argument is that since the object and effect of an active power is *ens factum,* a real product, and since no such power can operate on an ontologically deficient object, therefore, just as sight cannot see in the dark, God cannot do anything that goes against the very concept of being. But only one thing is contrary to the concept of being, namely, nonbeing ("nihil autem opponitur rationi entis, nisi non ens"). Hence, God cannot make that one and the same thing simultaneously is and is not ("hoc igitur Deus non potest, ut faciat simul unum et idem esse et non esse"). I do not see that apart from such ontotheological considerations, there is any reason for divorcing possibility from power.

From his logical understanding of modality Aquinas derives also the necessary fixity of the past—which not even God can undo.

> Deus non potest facere quod praeteritum non fuerit. Nam hoc etiam contradictionem includit: eiusdem namque necessitatis est aliquid esse dum est, et aliquid fuisse dum fuit.
>
> God cannot make that the past has not been. This too involves a contradiction: for indeed the same necessity pertains to something being while it is as to something having been while it has been.
>
> (*Summa c. Gent.* 2.25)

From our ordinary vantage point here and now the necessity of the past is, of course, so trivial that one does not care to mention it. What matters to us is the necessity of future events, or, more precisely, their necessitation by the present. The perceived success of the physical sciences in understanding it and bending it to our advantage is probably the chief source of the popular admiration for them. (And the popular contempt for philosophy would no doubt increase if more people knew that many of its practitioners—including some self-styled "realists"—dispute that any such necessitation exists). Our next—and final—task will be to examine the intellectual means employed by theoretical physics in the pursuit of that twofold aim. I shall maintain that they have consisted in grasping the seemingly blind fatality of events as a manifestation of conceptual relations. But we shall be able to see better how the necessity of natural processes can be epitomized by conceptual relations of a certain type if we first take a look at another, no less familiar, form of necessity.

When a speaker of English says that you *cannot* do a certain thing, he need not refer to an impossibility that constrains you with what I have been calling brutal necessity. For example, if you are told that in some parts of Antarctica you cannot sleep in the open air, this does not mean that you absolutely cannot do it, but that you cannot do it and still avoid freezing to death during the night. Should one rather say that you *can* but *may not* do it? However, ordinary usage stubbornly refuses to make a fast and clear distinction between 'can' and 'may'. Indeed, the former is the proper verb to use whenever a necessary connection is sensed—or assumed—between what is declared (conditionally) impossible and some result that one (absolutely) wants to preclude.

The same expression 'you cannot' is employed to convey the myriad impossibilities that entangle—and enable—our life in society. But the necessity here at play is plainly of another sort. And within this family, the conditional shades almost insensibly into the absolute. Consider a few examples:

(1) You cannot drive northwards on Fifth Avenue.

(2) You cannot read a Spanish text from right to left.

(3) You cannot obtain an Irish divorce.

(4) In most countries, you cannot sell real estate without a deed.

You can drive, of course, on a Sunday morning from the New York Public Library to Saint Patrick—you merely risk being stopped and fined by a policeman. Also, if you speak Spanish, you can manage, with a little exercise, to read fluently from right to left. Juan de la Cruz will sound like a Martian poet, but a poet all the same. However, his words will not make sense. Thus, whether the impossibility stated in (2) is absolute or conditional depends on what one means by reading. On the other hand, at the time of writing it is absolutely impossible to obtain a divorce in Ireland. An Irishman can indeed leave his Irish wife and never see her again. He can also divorce her in Mexico and marry a third party in Madagascar. But under Irish law they will still be married till death them do part. Example (4) illustrates the same idea, but there is an interesting twist to it, due to the mingling of physical with legal necessity. In exchange for a lump sum I can surrender my home in Puerto Rico to an illegal alien who does not wish to sign a legal document. But he will not own it. So he cannot sell it (without fraud). And he obviously cannot take it to another country where our verbal contract might be acknowledged as a valid sale.

Ever since Greek intellectuals pondered the contrast between φύσις and νόμος—nature and convention—there has been a philosophical tendency to lump into the latter category all forms of social necessity—linguistic, legal, ethical, etc.—and to oppose them to the former. Such polarity is evidently simplistic. In an attempt to defuse it, some 5th century Athenian coined the oxymoron ὁ νόμος τῆς φύσεως[16]—nature's convention—subsequently rendered in Latin as *lex naturae*, the law of nature. The phrase is still very much alive in papal encyclicals and epistemological tracts. We cannot delve here into its diverse and often conflicting connotations. We have to consider only a particular form of socially induced necessity which I believe can throw some light on the handling of natural necessity in physics; viz., the thoroughly conventional necessity that restricts the admissible moves in a game of strategy.

In every game there are things that a player can never do, or cannot do in certain situations. In sports, however, the impossibility is conditional: the inadmissible move—e.g., a forward pass in rugby football—cannot be undone, so if such a move is performed—and seen by the umpire—the culprit is penalized. But in pure games of strategy, such as chess or tic-tac-toe, impossible moves are absolutely impossible: you cannot make your queen jump over a bishop and still pretend that you are playing chess. The moves are best thought of as a succession of ideal elements, symbolized by certain physical events. Chess moves are represented by motions of the pieces on the chessboard, but not every physically possible motion of a piece represents a move in the game. At any given stage of the game, the person whose turn it is to play can choose her move from a fixed set characteristic of the position that was reached in the preceding move. (In some positions, she may be restricted to a single move.) The chosen move leads to another position, in which the rival player can in turn choose from a fixed set a move leading to another position, etc. Thus, the alternative courses that a game can follow from any given point are adequately conceived as paths in an oriented graph,[17] whose vertices are chess positions and whose edges are chess moves. Such a graph is a subgraph of what I shall call the graph of chess, an oriented graph that contains a vertex representing the initial position of the game and all the paths that issue from that position and can be followed by successively playing legal moves.[18] Let a complete path be one that begins at the initial position and ends in checkmate or in a legal draw. Every conceivable game of chess is represented by some initial segment of a complete path, and almost every complete path represents a legal game of chess. (The rare exceptions are explained towards the end of note 18). Each time two persons sit down to play chess, every game thus represented in the graph of chess is, in principle, possible: not, indeed, because in some world or other, counterparts of either player really play that game, but because *hic et nunc*, in this very world, the one and only in which both players are alive, their free submission to the rules of chess enables all those games. As their particular game progresses, more and more of the alternatives initially open become closed. Eventually, a position may be reached where, say, Black checkmates in four moves, i.e., a position such that by judiciously choosing his move on his next three turns Black can constrain White into a path ending with checkmate after White's third turn. As this case clearly shows, the possibilities and necessities in question are to be understood in the ordinary sense: they concern the players' actual powers at a particular juncture in their lives. Yet the graph of chess furnishes a universal scheme which covers many such junctures and furnishes an exact idea of the alternatives available at each one of them. The graph even provides the means of calculating the utility of the various possible moves for the purpose of winning the game. Thereby, the players' conduct of the game becomes understandable and—up to a point—predictable. Intelligibility and predictability follow here from the free agreement between two intelligent creatures to play a game of strategy they both conceive in the same way. But why shouldn't inanimate things likewise fall of themselves—or by the Demiurge's will—under a conceptual pattern spanning their evolution in time so that their past enables and constrains their future? Mathematical physics was originally rooted in just this belief. Whether it can long survive without it is indeed a moot question.

5.2 Geometry

Games of strategy have been around for a long time. Obviously they cannot be played without some form of awareness of their underlying structure. But a mathematical theory of games was not developed until fairly recently. The deliberate investigation of mathematical structures as a source of necessity and possibility did not begin with the study of games but with the study of geometry.

We tend to think of geometry as the science of space. Kant, who viewed it this way, saw space, in turn, as the "condition of possibility" of our perception of things; but it is perhaps more appropriate to think of it as a primary enabling and constraining condition of action—of our own action on things and of their reaction and interaction. To the Greek inventors of geometry, who did not even have a term equivalent to 'space', such talk would have seemed abstruse. Euclid's *Elements* are concerned with figures, for the most part with flat figures (on a plane). The construction of figures with certain properties was a matter of practical interest to carpenters, architects, surveyors, etc. At some point, however, members of the educated leisure class were fascinated by the subject and pursued it just for the fun of it, beyond any practical need.[19] Euclid assumes that one can execute certain elementary constructions, viz., join any pair of points by a straight line, extend a given straight beyond either endpoint, draw a circle with any point for a center and any straight segment for a radius. Thereupon he shows how to construct other figures by judiciously combining the elementary constructions and demonstrates certain features which the figures thus constructed will inevitably possess. As the book progresses the construction tasks recede, the demonstrations prevail unrivalled, and the illusion is fostered that one is reading, not about possibilities available here and now (and the constraints they are subject to), but about the properties and relations of timeless objects in another world. Demonstration follows upon demonstration in the "long chains of reasons" admired by Descartes.

Some geometrical arguments are so well knit that one may be tempted to think that the necessity displayed by them is purely verbal, that the conclusions are implicit in the very meaning of the words employed. Take, for instance, the striking fact that the three perpendicular bisectors of any triangle meet at one point. Let A, B, and C be the vertices of a triangle and let the perpendicular bisectors of AB and BC meet at Q. Since QA = QB = QC, Q evidently lies on the perpendicular bisector of the remaining side AC. But, of course, the expression 'perpendicular bisector of XY' does not designate the locus of points equidistant from points X and Y but rather the perpendicular through the midpoint of segment XY. That the latter constitutes the said locus is, however, so obvious that it may well go without saying. Just turn over the plane while holding the perpendicular bisector of XY fixed: X and Y will exchange places while every point on the perpendicular bisector stays put; hence, any such point lies at the same distance from X and Y. I have resorted here to an idea of rigid motion which can no doubt be verbalized and put forward as a definition of the subject matter of geometry. Necessary and sufficient conditions for every construction and demonstration in Euclid can also be stated in other ways.[20] But the bonds that such verbalizations disclose between the components of geometrical figures are not established by linguistic usage, let alone by linguistic agreement. Verbal necessity may link the premises to the conclusion in a geometrical proof but not the points and lines of the plane among themselves in the manner entailed by those premises.

The illustration I shall now give will, I hope, make my meaning clearer. Like the verbalizations I have alluded to, it involves the modern (Cartesian) understanding of the Euclidian plane. It also involves some—fairly simple—notions of complex analysis which, unfortunately, I cannot explain here. However, the following preliminaries should enable the reader unacquainted with those notions to see the drift of my argument. Other readers ought to proceed directly to the beginning of the next paragraph. Complex numbers were originally introduced to provide solutions for every quadratic equation. They can be readily conceived as ordered pairs of real numbers, subject to the following rules of addition and multiplication: $\langle a,b\rangle + \langle c,d\rangle = \langle a + c, b + d\rangle$; $\langle a,b\rangle \times \langle c,d\rangle = \langle ac - bd, ad + bc\rangle$. Endowed with these operations, the set $\{\langle a,b\rangle \mid a,b \in \mathbf{R}\}$ is a complete field, the field of complex numbers, usually denoted by **C**. Note that the said operations, restricted to $\{\langle a,0\rangle \mid a \in \mathbf{R}\}$, characterize a subfield of **C**, isomorphic to **R** itself. (Put $b = d = 0$ in the above definition of complex addition and multiplication.) Every complex number whose second term is equal to 0 may therefore be identified with the real number that is its first term, viz., $\langle a,0\rangle = a$; in particular, $\langle 1,0\rangle = 1$. Note further that $\langle 0,1\rangle \times \langle 0,1\rangle = \langle -1,0\rangle = -1$. Thus, $\pm\langle 0,1\rangle$ are the solutions of the quadratic equation $x^2 + 1 = 0$. The complex number $\langle 0,1\rangle$ is usually denoted by i. Evidently, every complex number can be written as a linear combination (with real coefficients) of 1 and i, thus: $\langle a,b\rangle = (\langle a,0\rangle \times \langle 1,0\rangle) + (\langle b,0\rangle \times \langle 0,1\rangle) = a + bi$. If we now assign Cartesian coordinates to the Euclidian plane, we can identify the complex number $\langle a,b\rangle$ with the point with abscissa a and ordinate b, or, better still, with the vector from the origin to that point. In this interpretation, the addition of complex numbers as defined above agrees with the standard vector addition (as displayed in the familiar rule of the "parallelogram of forces"). To interpret multiplication we go over to polar coordinates. Let ϕ be the angle formed in the positive, i.e., counterclockwise, sense between the axis of the abscissae and the vector we have identified with $\langle a,b\rangle$. Thus, $\phi = \tan^{-1}(a/b)$. Put $\rho = \sqrt{(a^2 + b^2)}$. Clearly, then, $\langle a,b\rangle = a + bi = \rho(\cos\phi + i\sin\phi)$. ϕ is the argument or amplitude of $\langle a,b\rangle$; ρ is its modulus or absolute value. If, as is usually done, we denote a complex number by a single letter, say, z, its argument is denoted by $\arg z$ and its modulus by $|z|$. A couple of trials should persuade the reader that multiplication of any complex number by another has the twofold effect of stretching (or contracting) the former by the latter's modulus and rotating it by its argument. End of preliminaries.

Consider the mappings of $\mathbf{C}$ into $\mathbf{C}$, i.e., in the given interpretation, the mappings of the Euclidian plane into itself. (Mappings and the relevant notation were explained in Section 2.8.3.) By a region I mean a non-empty open connected subset of the plane (e.g., the interior of a polygon). A mapping $f: \mathbf{C} \to \mathbf{C}$; $z \mapsto f(z)$ is said to be regular on a region M if $f(z)$ changes smoothly as z ranges over M. More precisely, f is regular at $\zeta \in M$ if there is a complex number $f'(\zeta)$ which meets the following condition: for every positive real number ε there is a disk δ, centered at ζ, such that every z in this disk satisfies the inequality $|((f(\zeta) - f(z))/(\zeta - z)) - f'(\zeta)| < \varepsilon$. f is regular on M if it is regular at every $\zeta \in M$. Suppose now that Γ is a closed path in M, surrounding a simply connected[21] subregion of M, which we shall call $I(\Gamma)$ (I for interior). Then, if f is regular on M, its value at any point $z \in I(\Gamma)$ is given by

$$f(z) = \frac{1}{2\pi i}\int_{\Gamma} \frac{f(z)}{\zeta - z}\,d\zeta$$

the line integral being taken in the positive (i.e., counterclockwise) sense over the path Γ. The reader need not worry about the definition and the existence of the integral. What matters for my purpose here is that according to the theorem just quoted, the values that *any* mapping $f: \mathbf{C} \to \mathbf{C}$ takes on the boundary of a simply connected region of the plane uniquely determine its value at each and every point inside that region, provided that both the latter and its boundary are included in a region on which f is regular. (Apart from this condition, f is *completely arbitrary.*) The theorem can, of course, be proved by standard deduction from the axioms that characterize the field **C** and the definitions of the words employed. But the bond it reveals between the points of the plane should not be classified as one of "verbal" or "logical" necessity; not, at any rate, if logical necessity shapes up as inference and we do not countenance inferences from uncountably many premises. I, for one, would judge it a hypertrophy of logic to view every value that the regular mapping f takes on the uncountably infinite points enclosed by the path Γ as the conclusion of an inference which has for premises the values of f on the uncountably infinite points of Γ.

Euclid did not spell out every condition on which his constructions and demonstrations depend. It is a monument to his mathematical insight that he did make explicit one seemingly out-of-the-way requirement—the notorious Postulate 5—without which his whole project of reconstructing geometrical figures by sequences of elementary constructions would founder. As Paul Lorenzen (1984) has aptly noted, the enterprise of geometry, as the Greeks understood it, presupposes the repeatability of such sequences with homologous results. Suppose, for instance, that on the strength of Euclid's Postulates 1–3, I pick out two points A and B, draw the segment AB, draw AC equal and perpendicular to AB, bisect angle CAB, mark on its bisector β the segment AH equal to AB, draw the straight λ through H perpendicular to β and coplanar with AB, and finally verify that λ meets the extensions of AB and AC beyond B and C. (The reader is advised to sketch this exercise on a piece of paper.) The Euclidian program would plainly make no sense unless every time that the same series of constructions is repeated, starting from *any* pair of points A and B, the same verification ensues, viz., that the perpendicular λ to the bisector β at the (arbitrarily chosen) distance AB from A eventually meets the two sides of the right angle whose vertex is at A. However, in this case, repeatability is guaranteed only if the requirement put forward in Euclid's fifth postulate is satisfied.[22] The gist of that requirement can be stated as follows: If two coplanar straights λ and μ form unequal pairs of internal angles with a transversal τ, λ and μ meet on the side of τ on which they form the smaller pair of internal angles.[23] Of course, if the two pairs of angles are only slightly different, the intersection of λ and μ is apt to lie very far away. Thus, the feasibility of the Euclidian program for the construction and analysis of finite figures turns on far-flung features of the indefinitely extended plane. There are conditions equivalent to Postulate 5 which do not openly speak of matters so remote. For example: (i) Given a triangle with angles α, β, γ, it is possible to construct another, smaller triangle with the same angles (Wallis 1693). Also: (ii) It is possible to construct one rectangle

(Saccheri 1733). These apparently innocuous demands seem more germane to Euclid's overall constructive, finitistic approach to geometry than the requirement that two coplanar straights *nearly* orthogonal to a common transversal should meet. We should thank and admire Euclid—or his anonymous source—for having singled out the latter, more obviously compromising condition for explicit formulation. By thus drawing attention to the unifying structure that embraces and supports all geometrical constructions, he laid the groundwork for the modern view of geometry as concerned with space.

The repeatability of geometrical constructions is not indifferent to experimental physics. The design and interpretation of physical experiments generally presupposes some geometrical articulation of their place of occurrence. The commonest and humblest precision instruments sport simple shapes which are critically relevant to their performance, and their manufacturers assume as a matter of course that the elementary Euclidian constructions can be repeated *ad libitum* in the terms described above. Because of this, Hugo Dingler maintained that physics is irrevocably committed to Euclid's Fifth Postulate and may not toy with non-Euclidian geometries. At most, the latter can furnish useful computational tools, with no claim to reality.[24] Like so many other philosophers of science, Dingler apparently forgot that in physics every quantitative concept is blurred, i.e., is predicated with an agreed margin of error. Thus, there is no inconsistency in setting up a relativistic model of the universe with nonzero spatial Riemannian curvature on the strength of data gathered with a telescope manufactured according to Euclidian specifications, for in a small neighborhood of the observatory the Riemannian metric of the model deviates insignificantly from that of the local tangent space. (Think of an architect who plans to have a piazza tiled with marble squares, although he knows very well that a surface levelled with the assistance of terrestrial gravity is not flat and therefore, strictly speaking, does not allow such tiling.) And, of course, the overall site of the experiment at whose receiving end our telescope stands can be articulated as a fragment of a curved spacetime. Let me illustrate this with an example from relativistic cosmology. As is well known, the light from distant galaxies is redshifted, i.e., its spectrum displays familiar emission and absorption lines at frequencies lower than the ones predicted for such lines by atomic theory and measured on their analogues in terrestrial laboratories. In relativistic cosmology the redshift in question is attributed to the "expansion of the universe" and is therefore indicative of the time elapsed since the light was emitted. To ascertain the age of a given extragalactic light signal by measuring its redshift, the relativistic astronomer may not simply project the Euclidian trajectory of the incoming rays and figure out its length in light-years but must conceive the signal's history—its "worldline"—as a path in a curved spacetime. I do not see anything wrong with this. Indeed, from this standpoint the very reading of the data depends on a conjectural estimate of the spacetime curvature. But that is just one more uncertainty that one must learn to live with in our uncertain world.[25]

On the other hand, there are geometric propositions which do not admit approximation but are either exactly true or exactly false, and if true, then inescapably so. We are all acquainted with knots that become untied and others that only become tighter when one pulls the ends of the string. If a given knot belongs to one of these two kinds, so will every copy of it, no matter when or where or from what material it is made. A knot's behavior follows inexorably from the manner in which the string is threaded. Mathematicians can exactly describe such threading and also derive from first principles the kind of knot it will produce. But surely the tightening of some knots and the untying of others are not just matters of verbal necessity. Likewise, howsoever you divide the surface of a ball into patches, there will always be a manner of painting them solid red, blue, green, and yellow so that any two patches sharing a borderline display different colors; but if instead of a ball you are dealing with a ring-shaped object, you may require up to six colors to do the trick. Such properties can, of course, be inferred from a suitable definition of ball-shapedness or ring-shapedness, but again, it is no mere matter of linguistic preference how you have to define these shapes in order to license such inferences.

5.3 Mathematical physics

Given two points *A* and *B*, the existence of an intersection *C* between any two straights through *A* and *B*, the size of the angles they make at *C*, and the distance from *C* to the segment *AB* are completely determined. So much follows from the conditions spelled out in the axioms of Euclidian geometry. They impose necessary connections between the parts and features of any constructible geometrical figure whereby they both enable and constrain the human activity of constructing it. The axioms tell us what such a figure must have looked like if it was drawn, what it will look like if it will be drawn. But they do not establish any links between different times, between, say, yesterday and tomorrow. This is usually expressed by saying that geometrical relations are timeless, meaning that they hold at each particular time.

The necessity we see and feel in nature, and which we would wish to fend off and turn to our advantage, is not just timebound but timebinding: it ties the future to the present and the past, forcing change and preordaining its outcome, channelling the flux of events. At first blush, it has precious little to do with geometry. Yet the amazing contribution of modern physics to the exploitation of natural necessity by man has resulted from understanding it on the analogy of geometric necessity, through the representation of natural processes by mathematical structures. This was made possible by one of the most remarkable feats of human thought: the conception of time—past, present, and future—as a single, homogeneous linear continuum.[26]

A lucid explication of what it takes to be a linear continuum was not available until the late 19th century;[27] but the science of motion established two centuries earlier by Newton clearly presupposes that time has the same structure as the trajectory of a free particle in space. Indeed, when Galileo, in the Third Day of the *Discorsi*, let a line represent the time in which a certain space is traversed by a body in uniformly accelerated motion, and used the geometrical properties of that line to establish a functional relation between travelled distances and travel times, he must have understood that the time can be mapped bijectively onto the line, so that each segment discernible in the latter uniquely corresponds to a distinct subinterval of the former (Galileo, EN, VIII, 208–10; cf. EN, VIII, 85. See Section 3.1). A similar understanding was already implicit in the use of geometrical methods in Greek astronomy and also in the use of kinetic methods in Greek geometry itself. The ancient mathematicians did not feel bound by Euclid's short list of elementary constructions and worked on figures that cannot be drawn with ruler and compass. Some of these figures are defined by the motion of a point under certain precise conditions. For instance, the spiral studied by Archimedes in *De lineis spiralibus* is drawn by a point moving with constant speed (ἰσοταχέως) from a point, along a straight line rotating with constant speed—on a plane—about that point. The construction assigns a definite point on the plane to each instant of the motion, viz., the position reached at that instant by the moving point. The geometrical properties of the spiral depend, of course, on this definition, so their demonstration must involve a comparison between spatial magnitudes—lengths, angles, areas—and times. To establish his arguments on solid conceptual grounds, Archimedes begins his treatise by proving the following lemma: If a point moves with constant speed along any line and two lengths are taken on it, the latter have the same ratio to one another as the times in which the point traversed them. Archimedes reasons thus: Let AB be the line along which the point moves, and take on it the segments ΓΔ and ΔE. Let the time in which the point traverses ΓΔ be ZH, and HΘ the time in which it traverses ΔE. "It must be shown that the segment ΓΔ has to the segment ΔE the same ratio that the time ZH has to ΗΘ" (δεικτέον, ὅτι τὸν αὐτὸν ἔχοντι λόγον ἁ ΓΔ γραμμὰ ποτὶ τὰν ΔΕ γραμμάν, ὅν ὁ χρόνος ΖΗ ποτὶ τὸν ΗΘ—*De lin. spir.*, I). The proof proceeds by taking equimultiples of each length and the corresponding time, in accordance with the theory of proportions explained in Euclid's Book V. It is tacitly assumed that a point moving with constant speed traverses equal distances in equal times. As Dijksterhuis remarks, "it may seem that the proposition follows at once from this." However, a proof is required to show that the proposition holds also when the segments ΓΔ and ΔΕ happen to be incommensurable (Dijksterhuis 1987, p. 141).[28]

When Archimedes wrote *On Spirals,* the practice of mapping time into space was at least a century old. It was already involved in the geometrical models of planetary motions devised by Eudoxus, Plato's contemporary and friend, who also created the theory of proportions employed by Archimedes in his proof. Eudoxus associated each planet with an n-tuple of concentric spheres, such that (i) their common center is the center of the Earth; (ii) the first sphere rotates about the poles in one day; (iii) the ith sphere rotates with a constant speed of its own about an axis fixed on the $(i - 1)$th sphere; (iv) the planet lies on the equator of the n-th sphere. If the number of the moving spheres and their respective axes and speeds of rotation are suitably assigned, the astronomer should be able to calculate from the planet's current position its declination and right ascension at any time. The description of a Eudoxian planetary model with given parameters can be read—on the analogy of the above definition of the Archimedean spiral—as the definition of a special kind of curve—e.g., a Venusian or a Martian trajectory—drawn by a point on any spherical surface, such as the vault of heaven or the ceiling of a planetarium. A body affixed to such a point instantiates the Eudoxian concept of a certain planet. The radius vector of the planet (i.e., the radius which carries—*vehit*—it on its tip) points in such-and-such directions at such-and-such times with the same inexorable necessity as the position vector of an Archimedean spiral reaches at prescribed times the third, the fourth, . . . the nth turn of the curve, or as any side of an equilateral triangle makes internal angles of $\pi/3$ radians with the other two. The future path of the bright spot we are now observing is fixed and timed by what it *is* at present, if indeed it *is*, say, a venus or a mars in the sense aforesaid. The success of Eudoxian astronomy never measured up to its true potential, chiefly, I suppose, due to the inaccuracy of naked eye observation and to the tremendous difficulty of estimating the appropriate parameters by manual computation (with zeroless arithmetic!). It was nevertheless sufficient to win Plato's support for the program of ascertaining "by what hypothetical uniform, ordered, and circular motions the phenomena regarding the motions of the so-called wandering stars can be preserved."[29]

Compared with Eudoxian astronomy, modern mathematical physics displays a pageant of astonishing richness. Thanks to the wonderful invention of mathematical analysis it is able to tie the present to the future and the past with "bonds of necessity" and to save the phenomena, even if their evolution is not reducible to a combination of uniform circular motions. Is the obvious difference between modern physics and ancient astronomy merely a matter of complexity—and predictive success? Or have Newton and Maxwell, Einstein and Dirac, reached for—and sometimes attained—an essentially deeper understanding of events than Eudoxus or Hipparchus? Much of the 20th century debate in the philosophy of science turns, openly or covertly, around this issue. For my own part, I am inclined to view the Eudoxian program as a paradigm of mathematical physics, *provided that*—as above—*it is interpreted realistically*.[30] There are, however, some important differences between it and the generally acknowledged Newtonian paradigm which deserve discussion. Let me mention briefly the three that come to my mind.

(1) Newton's physics is not a collection of (related) mathematical recipes for the prediction of regular occurrences in nature but a system of mathematical principles for natural philosophy. Its basic notions of time and space, mass and force, bound together in the Laws of Motion, are ultimately meant to account for *all* phenomena, not just for those observable in this or that segment of our environment.

(2) In all theories designed after the Newtonian paradigm, the states of physical systems are described and their evolution is explained in terms of purportedly universal properties of matter. Specifically, Newton's extraordinarily successful unified theory of planetary motion and free fall subsumed these two seemingly disparate families of physical processes by postulating a mutual attraction, dependent solely on mass and distance, between all pieces of matter. Second-generation Newtonians took this to mean that each piece of matter was inherently the seat of an attractive force, acting on every other piece of matter throughout the entire universe. Subsequently, physicists "discovered" further universal forces: electricity and magnetism (later unified by Ampère and, more successfully, by Maxwell), the so-called weak force manifested in radioactivity (unified with electromagnetism by Salam and Weinberg), and the strong force which keeps the positive electric charges in the atomic nucleus close together. Current research aims at unifying all acknowledged elementary forces in a single theory.

(3) In Newtonian physics the future and past states of a physical system are linked by differential equations to the forces actually working on it.

The program of universal physics was initiated ca. 1600 by Kepler and Galileo and was subsequently taken up by Descartes and his followers. It was deeply at variance with the two-tiered Aristotelian worldview favored by 16th century scholastics, but it could draw inspiration from Lucretius and other ancient writers, and ultimately from the Presocratics. Force was dismissed by the Cartesians as an obscure and indistinct idea, but it can, of course, be traced back to Greek philosophy and to prephilosophical common sense. Differential equations, on the other hand, though adumbrated by Galileo, are Newton's very own creation.[31]

Differential equations lie at the heart of mathematical physics and contain the key to the modern understanding of physical necessity. To see how they perform this role, I propose to explain the simplest form of differential equation and to briefly indicate (in notes) various ways of generalizing it. Let U be an arbitrary subset of $\mathbf{R}^{n+1}$. We consider a continuous mapping f of U into $\mathbf{R}^n$ by $\langle t, \mathbf{r}\rangle \mapsto f(t, \mathbf{r})$ $(t \in \mathbf{R}, \mathbf{r} \in \mathbf{R}^n)$. The following expression is then said to be an *ordinary differential equation of the first order.*

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = f(t,\mathbf{r}) \tag{1}$$

We shall say that f is κ-Lipschitzian in $\mathbf{r} \in \mathbf{R}^n$ for some given positive real number κ, if, whenever $\{\langle t, \mathbf{r}_1\rangle, \langle t, \mathbf{r}_2\rangle\} \subset U \subset \mathbf{R}^{n+1}$,

$$|f(t, \mathbf{r}_1) - f(t, \mathbf{r}_2)| \leq \kappa|\mathbf{r}_1 - \mathbf{r}_2| \tag{2}$$

An *exact solution* of eqn. (1) is any mapping $\varphi: \mathbf{I} \to \mathbf{R}^n$ (where $\mathbf{I}$ is an arbitrary interval in $\mathbf{R}$) that meets the following three conditions:

(i) The first derivative φ' is defined and continuous on the interior of $\mathbf{I}$.

(ii) For every $t \in \mathbf{I}$, $\langle t, \varphi(t)\rangle \in U$.

(iii) For every $t \in \mathbf{I}$, $\varphi'(t) = f(t, \varphi(t))$.

A mapping $\varphi: \mathbf{I} \to \mathbf{R}^n$ is said to be an ε-*approximate* solution of eqn. (1)—for some real number $\varepsilon > 0$—if it meets conditions (i) and (ii) but satisfies the following requirement instead of (iii):

(iiiε) For every $t \in \mathbf{I}$, $|\varphi'(t) - f(t, \varphi(t))| \leq \varepsilon$.

If the derivative φ' becomes discontinuous on a finite subset $\{t_1, \ldots, t_n\} \subset \mathbf{I}$, φ is said to be an ε-*approximate piecewise* solution of eqn. (1) if condition (iiiε) is satisfied everywhere on $\mathbf{I}\backslash\{t_1, \ldots, t_n\}$, and for every i $(1 \leq i \leq n)$ we have that $|\varphi'_<(t_i) - f(t_i, \varphi(t_i))| \leq \varepsilon$ and $|\varphi'_>(t_i) - f(t_i, \varphi(t_i))| \leq \varepsilon$ (where $\varphi'_<$ and $\varphi'_>$ denote the derivative "from the right" and "from the left," respectively).

Suppose now that the mapping f on the right-hand side of eqn. (1) is κ-Lipschitzian in $\mathbf{r} \in \mathbf{R}^n$, for a given $\kappa > 0$. It can be proved that if φ_1 and φ_2 are two exact solutions of eqn. (1) defined on the same interval $\mathbf{I} \subset \mathbf{R}$, such that for some $t_0 \in \mathbf{I}$, $\varphi_1(t_0) = \varphi_2(t_0)$, then φ_1 and φ_2 are identical. In other words, under the stated conditions, there is on any interval about a suitable point in $\mathbf{R}$ at most one exact solution of eqn. (1) that takes at that point a given value in $\mathbf{R}^n$.[32] Suppose, further, that the domain U of the said mapping f is a closed subset of $\mathbf{R}^{n+1}$ that contains the point $\langle t_0, \mathbf{r}_0 \rangle$, and that $\mathbf{I} \subset \mathbf{R}$ is a compact interval that contains t_0. It can be shown that, if for every $\varepsilon > 0$ there is an ε-approximate piecewise solution $\varphi_\varepsilon: \mathbf{I} \rightarrow \mathbf{R}^n$ of eqn. (1) such that $\varphi_\varepsilon(t_0) = \mathbf{r}_0$, there exists an exact solution $\varphi: \mathbf{I} \rightarrow \mathbf{R}^n$ of eqn. (1) such that $\varphi(t_0) = \mathbf{r}_0$. Evidently, this exact solution is unique (by the former theorem).[33]

As I mentioned earlier, the differential equations I have been discussing are those of the simplest kind. The concept of an ordinary differential equation of the first order can be readily generalized by substituting in the above definitions the complex number field $\mathbf{C}$ for the real number field $\mathbf{R}$, or by choosing the domain U of f in eqn. (1) to be a subset of $\mathbf{R} \times \mathbf{A}$ or of $\mathbf{C} \times \mathbf{B}$—where $\mathbf{A}$ is a real and $\mathbf{B}$ a complex Banach space (see Section 2.8.4). Ordinary differential equations of the nth order involve derivatives of their solutions of order up to n.[34] Partial differential equations have solutions defined on $\mathbf{R}^m$ or on $\mathbf{C}^m$ (for some integer $m > 1$) and involve their partial derivatives.[35] Theorems on the existence and uniqueness of solutions to differential equations of these further kinds have been demonstrated under diverse restrictive assumptions. It is due to the existence and uniqueness of solutions that under suitable circumstances, differential equations provide an unequalled grasp of physical necessity.

To see this, let us go back to the simple case of eqn. (1). Suppose that $\mathbf{r}$ represents a physical quantity and that eqn. (1) expresses its time rate of change. (Remember that $\mathbf{r} \in \mathbf{R}^n$; each value of $\mathbf{r}$ is therefore a list of n real numbers and may therefore represent the distances of n particles from a given point, or the three position coordinates and three velocity components of $n/6$ particles, or the position coordinates, velocity and acceleration components, masses, and electric charges of $n/11$ particles, etc.) Suppose, moreover, that the assumptions for the existence and uniqueness of solutions are satisfied on a domain $U \subset \mathbf{R}^{n+1}$. Then, if the value of $\mathbf{r}$ is given at some time $t_0 \in \pi_1(U)$, it is thereby fixed for every time $t \in \pi_1(U)$. If the stated conditions hold, nothing—not even the Will of God—can change the course of the physical quantity represented by $\mathbf{r}$. Thus, if a physical process can be adequately conceived as the evolution in time of a quantity $\mathbf{r}$ governed by the differential equation (1) its necessity will thereby be finally grasped and understood. The same can be said, mutatis mutandis, if the adequate conception of a physical process involves a more general form of differential equation, provided, of course, that the appropriate assumptions for the existence and uniqueness of solutions are fulfilled. Note that this manner of understanding converts natural necessity into mathematical, and thus conceptual, necessity.[36] It stands to reason that some such conversion is required for an *understanding* of necessity.

The following example should further clarify the power of the mathematical physicist's approach to natural necessity, as well as its limitations. In the Hamiltonian formulation of Classical Mechanics, the evolution of an isolated system of n particles is fully determined by the current values of $6n$ functions of the time. They may be chosen to stand for the $3n$ position coordinates and the $3n$ momentum components of the particles in Euclidian space. Since we take the mechanical system to be isolated, its total energy H depends exclusively on the current position and momenta. If t denotes time and q_{hk} and p_{hk} denote respectively the hth position coordinate and the hth momentum component of the kth particle, the evolution of the mechanical system is governed by the following system of $6n$ partial differential equations of the first order:

$$\frac{\mathrm{d}q_{hk}}{\mathrm{d}t} = \frac{\partial H}{\partial p_{hk}} \qquad \frac{\mathrm{d}p_{hk}}{\mathrm{d}t} = -\frac{\partial H}{\partial q_{hk}} \qquad (h = 1, 2, 3;\ k = 1, \ldots, n) \tag{3}$$

A solution of eqns. (3) is a smooth curve in $\mathbf{R}^{6n}$ (or $6n$-dimensional Euclidian space), each point of which encodes the position coordinates $q_{hk}(t)$ and the momentum components $p_{hk}(t)$ of the n particles at a particular time t. If the conditions for the existence and uniqueness of solutions are fulfilled, there is one and only one such curve through any given point in $\mathbf{R}^{6n}$. Obviously, in that case, the position and momenta of the particles at any given time fully determine the evolution of the mechanical system before and after that time. Enraptured by this vision, Laplace stated his famous claim:

> An intelligence that knew, for a given instant, all the forces acting in nature, as well as the positions of all the things that constitute it, and who was capable of subjecting these data to analysis, would embrace in a single formula the motions of the largest bodies and those of the lightest atom. For her nothing would be uncertain, and the future, like the past, would be present to her eyes.
>
> (Laplace, OC, vol. VIII, pp. vi–vii)[37]

But even Laplace would acknowledge that human science, though "incessantly approaching" the intelligence just described, "will always remain infinitely far from it."[38] Indeed, Hamilton's equations hold in the above form (3) for an isolated system, but no physical system we will ever come across is really isolated. We do our best to seclude our experiments and our machinery from the rest of the world, at least insofar as the quantities of interest to us are concerned. But we never succeed perfectly or permanently. The Solar System lies so far away from all other significant gravitational sources that it provides a virtually insuperable paradigm of Laplacian determinism. It would, however, come apart and all predictions regarding it would be invalidated if the Sun came close to another star. Moreover, as we know, all the data of physics are more or less blurred, and of course blurred values will not determine unique solutions when substituted in differential equations such as (1) or (3). Based on this consideration, Popper (1950) has argued that contrary to the vulgar opinion, determinism is alien to classical physics. But even if one disregards the effects of blurredness, determinism encounters severe limitations in both classical and contemporary physics, as John Earman has shown in his *Primer on Determinism* (1986). I cannot summarize here this admirable book, which no budding philosopher of nature and natural science should go without. Let me just hurriedly say that universal determinism fails in a Newtonian world chiefly because there is no upper bound to the speed of signals. As a consequence of this, the spacetime region over which the values of physical quantities can be computed with certainty from an accurate and exhaustive knowledge of their values at a particular time is confined to that time alone. This limitation is overcome by Special Relativity, which offers an intellectual oasis to determinists who are ready to assume a priori that they live in true Minkowski spacetime, without missing points or other topological anomalies. However, as Einstein was quick to see, gravitational phenomena do not fit comfortably in a Minkowski spacetime.[39] The Theory of General Relativity developed by Einstein to cope with gravity provided, for the first time in history, a mathematical framework in which the deterministic evolution of the universe can be roughly yet plausible represented. But this framework admits some pretty wild spacetime topologies. Worse still, under very general assumptions, a model of General Relativity typically contains so-called singularities, which, in turn, often imply a radical breakdown of determinism.[40] Finally, with the introduction of Quantum Mechanics in the 1920s physical determinism took a surprising twist. A solution of the Schrödinger differential equation does not describe the evolution of any quantity or set of quantities one might properly be said to observe but rather that of a mathematical object from whose successive values one can calculate the current *probability* of obtaining each of the admissible values of any specific observable quantity, should one choose to measure it. But such limitations of determinism cannot, in my view, detract from the significance of mathematical physics, or its success. For its aim is, of course, to understand natural necessity where we find it, not to yield some Procrustean scheme for conceiving everything as the outcome of necessity.

5.4 Cause and law

In the following I shall often refer generically to physical systems that are conceived, to within some suitable approximation, as evolving in time according to a law expressible by a differential equation, in the manner roughly indicated in Section 5.3. Let me call them GDE-systems. I shall also speak of GDE-processes, or GDE-descriptions, etc. 'GDE' may be read as 'governed by a differential equation', provided that this expression is taken in the precise sense sketched in that section. 'DE', as in 'DE-solution', should be understood to refer to the pertinent differential equation.

There is a tendency in philosophy to regard the several states in a GDE-process as constituting a causal chain in which any two successive states are linked to one another as cause and effect. That tendency would breed no trouble if those who follow it refrained from employing the term 'cause' in its ordinary prescientific sense. Such abstinence, however, would thwart their aim, which is to present the differential equations of mathematical physics as the appropriate means for intellectually grasping causal connections in nature, and to suggest that every genuine cause-effect pair is actually embedded in some GDE-process and that the world itself is just one big connected GDE-system.[41] The tendency in question has been no doubt assisted by the aura of imprecision surrounding the ordinary notion of 'cause'. In turn, it should be held responsible for much of the obscurity and uncertainty of traditional philosophical analyses of causation. It is only recently that a few sharp-sighted philosophers have sought to break the timeworn association between cause and law (Ducasse 1924, 1969; Anscombe 1971; Cartwright 1983), but other writers still struggle to reinforce it (e.g., Tooley 1987).[42] A proper discussion of the issue could fill a whole book, but it is well that I bring this one to a close with some remarks on the difference between the commonsense idea of causation and the scientific notion of GDE-evolution, and on the complementary roles which both modes of understanding play in physics.

Let me note, first of all, that in the sentence scheme

$$\text{‘}x \text{ causes } y\text{’} \qquad (1)$$

one cannot, without doing violence to the English language, fill the blanks with definite descriptions of two successive momentary states of a GDE-system. As D. Gasking once observed, it would be a most unnatural and strained use of the word 'cause' to say that the movement of a falling body at 64 feet per second two seconds after it was let go *is caused* by its moving at 32 feet per second one second earlier (1955, p. 480). What goes best into the subject place marked in (1) by *x* are names (or definite descriptions) of persons or animals; though obviously one can also fill it with the name of a reputedly active thing—the Sun, the sea—or process—a hurricane, an earthquake.[43] Note that these are just the sort of natural entities that prescientific man was wont to personify. You may, no doubt, comfortably say in English that such and such an effect has been caused or can be caused by a given *machine.* Indeed one might say that is just the sort of thing machines are for. But then, perhaps this is only another way of saying that one expects machines to do—with improved efficacy—the work earlier done by men and animals.

Such prescientific but still wholly current and by no means archaic use of the verb 'to cause' is consonant with the original meaning of the Greek noun αἰτία, which the Romans rendered as *causa,*[44] whence the English 'cause'. Before being taken up by philosophy, αἰτία meant 'responsibility', mostly in a bad sense, i.e., 'blame', 'guilt', and was also used, metonymically, to designate whoever or whatever was to blame for something.[45] Inquiry into the responsibility for a given state of affairs was not confined, however, to the persons or personifiable things who authored it, but it sought to single out any abstract feature by virtue of which they were in a position to do so. Thus Democritus, in the earliest documented use of the word by a philosopher, says that "ignorance of what is better is the cause of wrongdoing" (ἁμαρτίης αἰτίη ἡ ἀμαθίη τοῦ κρέσσονος—Diels-Kranz 68.B.83). Earlier on, Herodotus had set out to record the great deeds of the Greeks and the barbarians, "as well as the cause why they warred against each other" (καὶ δι' ἣν αἰτίην ἐπολέμησαν ἀλλήλλοισι—I, 1). He also wondered "what on earth was the cause that necessitated" (τὸ αἴτιον ὅ τι κοτὲ ἦν . . . τὸ ἀναγκάζον) Thracian lions to exclusively attack camels in Xerxes' camp, although they had never before "seen or experienced the beast" (VII, 125). About the same time the Hippocratic physicians were already looking for factors of every sort—climatic, nutritional, behavioral—that could be blamed for illness and pain. It is not surprising, therefore, that when Plato's Socrates tells his young friends about his short-lived affair with natural science, the term αἰτία should have pride of place:

> ὑπερήφανος γάρ μοι ἐδόκει εἶναι, εἰδέναι τὰς αἰτίας ἑκάστου, διὰ τί γίγνεται ἕκαστον καὶ διὰ τί ἀπόλλυται καὶ διὰ τί ἔστιν.

> It seemed splendid to me to know the causes of each thing: what is that by virtue of which each thing is born, and that by virtue of which it is destroyed, and that by virtue of which it is.
>
> (*Phaedo*, 96a)

It was for dealing with just these questions that Aristotle developed his doctrine of the four types of "cause": the matter, the form, the goal, and the agent (*Phys.* II, 3). This involves such a colossal expansion of the denotation of the word αἰτία that a recent commentator has suggested that it might be rendered better as 'explanatory factor'.[46] I do not commend such ploys, which merely serve to protect young readers from becoming aware of the plasticity of human thought. But I can see, of course, that of all four Aristotelian "causes" only the agent or "source of the change" (ἀρχή τῆς κινήσεως) falls under the pre-Aristotelian meaning of αἰτία and the ordinary acceptation of 'cause'.[47] Now, Aristotle's examples of an agent are usually men: the sculptor, the builder. He does, however, emphasize that it is the builder at work as such, *the builder building* (ὁ οἰκοδόμος οἰκοδομῶν), who causes the house to rise. This might seem to adumbrate the modern philosophical view that only events can properly count as causes. But Aristotle would have been baffled by the very idea that an event—let alone an instantaneous event—might *act* to *effect* another event. It is clear to him that to bring something about, to do it or make it, is the act or work (ἐνέργεια) of a continuant.[48]

The experience of authorship, of doing a deed, is one of the primary constituents of our self-awareness. A philosopher who pretends he has no notion of it will only succeed in *literally* making a fool of himself. Nor can he hide the fact that our ordinary concept of causation is intimately bound to that notion, either by being formed as a natural extension of it, or—what seems to me more likely—by functioning from the outset as the broader category under which our own doings fall. But neither the notion of authorship nor the concept of causation of which it is a paradigm case involves the idea of law. To understand that I am doing what I am doing does not in any way imply that there is a general rule connecting my (type of) activity with its (kind of) results. Of course, in settling questions of responsibility—e.g., in a criminal trial—it is useful, nay indispensable, to consider what sort of causes usually bring about, under like circumstances, like effects. But regularity is here an indicator of causality, not a defining trait of the causal relation itself. If one clearly and distinctly sees a man pierce a hole through someone else's skull with a single blow of his fist, one understands that he made the hole, although the occurrence is unprecedented and, one hopes, will not be repeated. In fact, if succession according to a law were implicit in the very concept of causation, narrative literature—which usually employs multiple varieties of the latter without supplying even an inkling of the former—would be, for the most part, unintelligible.

But even if the extension of '*x* causes *y*' were broadened to allow physical states as causes, one big difference would still separate the relation between cause and effect, as it is ordinarily understood, from the relation between two successive states of a GDE-system: the latter relation is symmetric, while the former is not. If state σ_1 at time τ_1 and state σ_2 at time τ_2 both lie on the unique DE-solution through each, σ_1 determines σ_2 no less than σ_2 determines σ_1. Indeed, knowledge of a particular state in such a GDE-evolution can serve equally well to predict later states and to retrodict earlier ones. Causation, on the other hand, is inherently asymmetric: the stallion is not begotten by its foal; the food does not cook the fire.[49]

To my mind, however, the main discrepancy between ordinary causal thinking or etiology and the nomology of differential equations lies in that the former thrives on discreteness, while the latter assumes continuity.[50] The analysis of a causal chain is not completed until every link in it has been pinpointed in the sequence it forms with its direct and indirect neighbors. But in the differentiable manifolds in which GDE-trajectories are embedded such full pinpointing is not even conceivable. A state in a GDE-evolution does not have an immediate successor through which its power of determination is, so to speak, passed forward in the causal chain as the torch in a relay race. If two terms in a causal chain are not directly connected, one is bound to ask for the intermediate links. On the other hand, if σ_1 and σ_2 are any two states on a GDE-trajectory, they are certainly not contiguous, yet the question "What precisely does σ_1 act on, in order eventually—after uncountably many like actions—to determine σ_2?" does not make any sense. In a GDE-evolution one state determines the others not by dint of a state-to-state transmission of efficacy but formally (so to speak), through the structure in which they are all comprised.[51]

Ordinary causal thinking is man-centered and context dependent. How it singles out effects and what it pinpoints as their causes will depend in each case on human uses and views and on the purpose at hand. Pragmatic considerations also guide the choice of a specific GDE-"model" to represent a fragment of reality, but within such a "model" relations are settled once and for all, and do not vary with perspective and interest. This fosters the impression that nomological thinking comes closer to the Truth than ordinary causal thinking, and that the latter is a remnant of primitive animism which should eventually disappear. However, it does not seem likely—or even conceivable—that this could ever happen. As a matter of fact, in real life these two disparate ways of understanding work fairly well together. Thus, the best way of securing an effect E is to find a GDE-process leading to E from initial or boundary conditions C which one is able to set up. In such a case, it is appropriate to say that E is caused by C or by the act A of setting up C or by the person P who executes A. But this does not entail that the succession of nondenumerably many GDE-events joining C to E is being regarded, per impossibile, as a causal chain: in the analysis proposed, the said succession is the single link between cause C and effect E. At times, a more detailed causal analysis may be useful. Consider, for instance, a particular computer input I. Shortly after I is entered, the corresponding output O turns up on the computer screen. From a "hardware" viewpoint, entering I causes O through a continuous electrodynamic GDE-process. But from the more familiar and superficially more perspicuous "software" viewpoint, I leads to O through a finite chain of conditional commands, determined by the input itself, the computer's initial state, and the computer program, say, $I = C_0 \rightarrow C_1 \rightarrow \ldots \rightarrow C_n \rightarrow O$. Obviously, one may properly say that the execution of C_i has led to, brought about, or, indeed, caused the execution of C_{i+1} $(0 \leq i \leq n)$. By this manner of speaking one in effect singles out n computer states besides I and O, and decomposes the single GDE-process between I and O into $n + 1$ stages linking those intermediate states. Such causal analysis can certainly assist one in ascertaining what changes in the input I may be required to obtain an output different from O, but it will not contribute (except perhaps heuristically) to understanding the physical process through which, given I, the system necessarily outputs O. Causal thinking is likewise inevitable in laboratory life. Our continued quest for more detailed and accurate GDE-representations of natural phenomena depends on experiments, which must, of course, be initiated and taken notice of by persons. An experiment to test our proffered nomological understanding of some kind of phenomenon will not make sense as such if its very occurrence is in turn understood nomologically, viz., as a minor event in a major GDE-development comprising, say, the whole history of the Galaxy. Such an experiment is performed on what is purportedly a GDE-system of the sort under study. This is kept as isolated as it is humanly possible, at least insofar as the relevant physical quantities are concerned. Yet the experiment must have been contrived by men and yield results they can take stock of. The experiment's double interface appears in the nomological representation as the initial or boundary conditions and the final state of the GDE-system (cf. the passage by Margenau, quoted in note 41). But in scientific practice, it must also be understood causally, as that which the experimenter does, and as what he achieves by doing it. Thus, at the heart of physics, as in other walks of life, we resort to more than just one mode of thought. The experimental interface must be conceived from either side of it in radically different ways. A frank recognition of the irreducible difference between etiology and nomology and of their complementary roles in physics will serve us better, I dare say, than the familiar philosophical attempts at refurbishing the former to make it look like the latter.[52]

PREFACE

ὁ τοιοῦτος νοῦς . . . εἴη ἄν καὶ ταύτῃ ποιητικός, ᾗ αὐτὸς αἴτιος τοῦ εἶναι πᾶσι τοῖς νοουμένοις.

Moreover, such an understanding is creative inasmuch as it brings every notion into being.

ALEXANDER OF APHRODISIAS†

The understanding of natural phenomena developed by mathematical physics since the 17th century stands today as one of the most remarkable achievements of human history. Rightly or wrongly, many regard it as a paradigm for every scientific endeavor. And its practical applications have, for good or ill, drastically altered the fabric of life.

In a lecture on the method of theoretical physics delivered at Oxford in 1933, Albert Einstein noted that the concepts and fundamental laws of physics are not derived by abstraction from experience, nor can they be justified by appealing to the nature of human reason, for they are "free inventions of the human mind" (Einstein 1934, p. 180). Clearly then, if we take our cue from him, the understanding wrought from those concepts and laws must be termed 'inventive' or 'creative'.

The aim of this book is to elucidate, at least in some important respects, the workings of that creative understanding. No attempt is made to unveil the mysteries surrounding intellectual creativity. But through the examination and interpretation of examples from the history of science, and the critical discussion of good and not so good ideas from recent philosophical literature, it seeks to throw light on the means and ends of the intellectual enterprise of physics.

The book is divided into five chapters, labelled "Observation," "Concepts," "Theories," "Probability," and "Necessity." The conventional starting-point with observation was chosen, not to pay lip service to the philosophy of inductivism, but to underscore its inadequacy. Observation without understanding is blind. We must grasp phenomena under universal concepts in order to make them out, and so make them into facts. This Kantian thesis is rehearsed and illustrated in Chapter 1. A strong argument for it follows from the preferential status of instrumental observation in science. The recorded modification of an instrument can yield information about the state of the observed object only to the extent that the latter is a necessary condition of the former. However, a necessary connection between them cannot be found by inspecting the instrument's dial, but must be read into it in the light of an overall understanding of the physical situation. In modern physics, the required understanding is supplied by the richly articulated grasp of physical systems afforded by physical theories.

Physical theories are in effect the unifying theme of the remaining four chapters. The formal examination of their typical (idealized) structure and mutual relations in Chapter 3 is preceded and prepared by the informal presentation and discussion, in Chapter 2, of a problem that has much exercised philosophers of science since the publication of Thomas S. Kuhn's influential essay on *The Structure of Scientific Revolutions* in 1962; namely, the alleged incomparability or, as the saying goes, "incommensurability" of succeeding theories. The problem arises if the identification of the objects of scientific discourse depends on the concepts employed for describing them, and there is no fixed set of concepts shared by all scientific theories. After discussing and dismissing with familiar arguments the view that such a fixed set of concepts is available, I embark on a lengthy criticism of Hilary Putnam's doctrine of reference without sense, according to which one can single out a determinate physical object, e.g., a physical magnitude, independently of how one conceives it. Although Putnam himself has discarded this doctrine, it is still favored by several authors.

My own approach to the problem of the incommensurability of physical theories is explained and defended in Chapter 2, especially in Sections 2.5 and 2.7. The problem can be posed in earnest only with respect to the theories of fundamental physics, which can, in turn, supply the requisite conceptual bridges between successive theories of narrower scope. The problem does not arise when a new theory of fundamental physics is reached—as Special Relativity was by Einstein in 1905—through internal criticism of the preceding theory. Yet even when that is not quite the case, a shared tradition of mathematical thought makes it possible to read the earlier theory in terms of the new one, or to devise an ad hoc common framework for the description and assessment of experimental data. But apart from such internal means of comparison, the fundamental theories of mathematical physics communicate across the common ground of understanding from which they grow and which they serve: the loose, unpretentious grasp of things and events in everyday life. This is not to say that science is accountable to common sense, or that scientific discourse should be translatable into ordinary language. But since the latter has not been—and presumably never will be—replaced by the former, it continues to supply the murky global perspective within which each particular theory of physics discerns the facet it seeks to conceive with clarity and precision.

In Chapter 3 I take the view that a physical theory conceives an open-ended host of physical situations as instances of a mathematical concept. This view was put forward by Joseph Sneed in *The Logical Structure of Mathematical Physics* (1971), and has recently been presented in a clearer and more elaborate form by Balzer, Moulines, and Sneed in *An Architectonic for Science* (1987). Sneed and his collaborators explicate the mathematical concept at the heart of a physical theory in set-theoretical terms, as a Bourbakian species of structure. This approach is also adopted here, not due to any sympathy for Bourbaki's ideas, but because they are good enough for the present job, and after several decades of dominance over the teaching of mathematics, they have become fairly well known. Sneed's analysis of the structure of a physical theory is illustrated and motivated with some examples from history and then formally explained. There follows some criticism of important aspects of Sneed's doctrine. It is shown that his distinction between the models and the potential models of a physical theory, although useful for explicating the familiar contrast between the *concepts* of physics and its *laws*, is in each case relative to the peculiar way chosen for reconstructing the central concept of the theory in question as a Bourbakian species of structure. Such relativity undermines Sneed's use of so-called partial potential models in the solution of his problem of theoretical terms. This, however, is a pseudoproblem, stemming from a refusal to countenance a genuinely creative understanding of natural phenomena, and nothing is lost by forgoing its purported solution.

Sneed and his associates have developed the means of conceiving both the links that can be established between several closely related or widely divergent theories and the constraints which bind together the different applications of a single theory. They readily account for the fact that the number and variety of applications of a physical theory can change without prejudice to its conceptual identity. But they do not tell us how a fragment or aspect of experience is turned into an application of a physical theory; how scientific thinking takes hold of a domain of reality, fills in the intelligibility gaps, and articulates it as a domain of objectivity. In Section 3.5 this problem is elucidated in the light of the work of Günther Ludwig. Ludwig has also originated the best available formal treatment of approximation in physics. His contribution and its implications for the coexistence and joint employment of seemingly incompatible theories are explained in Section 3.6. Sections 3.7 and 3.8 methodically examine the different types of relations that exist between physical theories and contain the book's last word on the problem of Chapter 2.

In Chapter 4 I interrupt the discussion of physical theories in general, to grapple with a difficult but rewarding illustration: the physicomathematical concept of probability or, more precisely, of a probability space. The choice of this example may seem questionable, for the species of structure *probability space* is not by itself the central mathematical concept of any physical theory. It does, however, occur as a constituent in several such concepts (just as the *real number field* occurs in all), and due to its comparative simplicity, it lends itself better for the didactic purpose of this chapter than the conceptual core of a full-fledged theory. A deeper motivation for choosing it lies in the controversies that surround it. What sort of physical reality is represented in physics by the concept of a probability space? There is no general agreement on this point, and a powerful school professes that there is no such reality at all. The issue had to be discussed before dealing with physical necessity and determinism in the final chapter. It is a welcome opportunity for considering the creative understanding at work as a source of objectivity.

That necessary connections between physical events are involved in the current understanding and utilization of scientific observations was the main result of Chapter 1. Chapter 5 shows how physical theories use mathematical concepts to embed natural phenomena in a tissue of such connections. The conception of a physical theory as a species of structure modelled by the theory's applications finds its ultimate vindication here. The chapter also includes some reflections on the sources of the notion of physical necessity in ordinary human experience (§5.1) and on the uneasy coexistence of physicomathematical determinism with commonsense causality (§5.4).

*****

The notes serve several purposes. A few of them explain technical terms, especially from mathematics. Others pursue special questions or try to ward off possible objections. Others provide illustrative quotations, mention sources, or make suggestions for further reading. I suppose that the main text can be understood without referring to the notes, but I expect that some readers will find them useful.

References to the literature are usually identified by the author's name followed by the year of publication. Exceptionally, when I refer to an edition published much later than the original work, I substitute a word or a few identifying letters for the year of publication. (The sole purpose of this is to avoid vexing anachronisms such as Einstein 1987 or Heraclitus 1855). Titles are given in References, at the end of the book. That list contains only works which have been mentioned in the main text or the notes and does not fully reflect my debt to other writers. I have tried my best to indicate the provenance of my ideas, but I do not always remember.

I have indulged in a minor deviation from standard English usage. I let the pronoun *she* stand for the noun *person* when the latter refers to an indeterminate human being. I see this as a tame—and etymologically justifiable—gesture against linguistic *machismo*. Of course, in my own language a person—*una persona*—is always referred to by the feminine third person pronoun *ella* even if she happens to be a male.