Cambridge University Press
978-0-521-45597-8 - The Observational Foundations of Physics
Alan Cook
Excerpt
More information

1

Introduction

1.1 The questions

MY PURPOSE IN THIS BOOK is practical and empirical, it is to attempt to unravel some ways in which the practice of physics determines the form and content of physics and physical theory. That is no novel undertaking. Eddington (1953) in particular claimed to derive many fundamental features of physics from deep epistemological principles, but as is well known, few have understood what he was about and fewer still have agreed that he was successful. My aim is less ambitious than that or of some philosophical discussions, it is to look at what physicists actually do in making observations and assessing their reliability, and to follow through the consequences of those practices for the theoretical structures of physics. Modest though that may seem, we shall find that it leads us into quite deep and intractable questions concerning the status of observation, the basis of inference and the reliability of physical knowledge.

A very striking feature of the physical sciences is that they are remarkably effective at predicting from past phenomena the nature of events yet to take place. Why should physics be so effective, and what does that tell us about the world of physics and our ways of gaining knowledge of it? The question has become the more acute

Cambridge University Press
978-0-521-45597-8 - The Observational Foundations of Physics
Alan Cook
Excerpt
More information

as it is realised that much of the behaviour of the natural world is at bottom chaotic, in the sense that conditions cannot be stated precisely enough for consequences to be predicted.

Most of this book is about physics in a rather restricted way and only at the end do I take up epistemological questions such as those at which I have just hinted, but there is one philosophical issue that must be faced at the beginning and then put aside until the last chapter. That is the question of the existence of a physical world independent of us, or more strictly, independent of me. Is there a real world that exists independently of whether I or anyone else is looking at it, or are all the ideas I have about a world external to me just the construction of my mind? My own opinion is that no answer can be given to that question. Either the realist position or the extreme idealist position (solipsism) can be the basis for a consistent account of what goes on in my mind, although I consider that it is difficult to hold any consistent intermediate position. In this book I write as a realist. In the first place, it is far more straightforward to do so than to write consistently as a solipsist. More importantly, and in the spirit of the overall approach I adopt, I believe that almost all physicists when working at the bench or with their pencil and paper or computers, behave as if there were a real world that will continue to exist whether or not they observe it or think about it (see d'Espagnat, 1989). Thus Pickering (1989) and Gooding, Pinch and Schaffer (1989) in their respective discussions of the *Uses of Experiment* explicitly accept that a real material world exists, and most of this book is written unquestioningly in that belief, but I shall return to it with other metaphysical matters at the end.

1.2 The nature of observation and of theory

While most physicists, so I think, pursue their vocation accepting the existence of a real world, independent of themselves, out there to be investigated, few are so naïve as to think that their observations give them direct unadulterated knowledge of that world. The formalism of quantum mechanics expresses the idea that our observations are the results of interactions between the world independent of us and the process of observing it, and the development of quantum mechanics has led scientists in other

Cambridge University Press
978-0-521-45597-8 - The Observational Foundations of Physics
Alan Cook
Excerpt
More information

fields also to appreciate that the results of experiment and observation depend on how we interact with the outside world when we perform those experiments and observations. I should comment here that I make no distinction between *experiment* and *observation*. Experiment commonly implies a more active approach to nature on the part of the observer, while observation is usually considered to be more passive. Those distinctions are irrelevant to the argument of this book.

There are two aspects to that dependence of the results of observation on our interaction with the outside world, an objective aspect and a subjective one, or, to put it slightly differently, a dependence that is the consequence of the physics and independent of the observer, and a dependence that follows from the personal competence or choices of the observer. The dependence incorporated in quantum mechanics is objective – it is expressed by operators of definite mathematical form, human factors do not come in. The dependence that comes from the design of experiment or technical competence is peculiar to the people doing the experiments. Social influences come in here. Without going all the way with sociologists of science who sometimes seem to imply that our view of nature has nothing objective to it at all but is entirely a social construct, or with literary theorists who would have us believe there is nothing beyond a text, it is still possible to recognise that the subjects on which physicists work and the ways in which they approach them, are certainly influenced by communal behaviour, although not wholly determined by it. Ziman (1978) asserted that the claim of science to be objective depends upon its being a social construct, created cooperatively, it is in his words, in the noetic domain, and d'Espagnat (1989) has made an important distinction between the subjective knowledge of an individual and the subjective elements in knowledge, common to a large group of individuals, on which as scientists they all agree. I am concerned in this book with that communally accepted body of knowledge.

Physics is supposed to be empirical and contingent, with observation primary and theory secondary but many philosophers of science have questioned that position. Hesse (1974) argued in some detail that even apparently simple observations depend on some theory for the interpretation of the raw response of an

Cambridge University Press
978-0-521-45597-8 - The Observational Foundations of Physics
Alan Cook
Excerpt
More information

instrument, and her argument could be put more forcibly today (Hacking, 1983). Consider for example, the measurement of the intensity of light emitted by a black body. The light falls on a semiconductor detector that generates an electrical signal that causes another semiconductor device, a digital voltmeter, to emit a train of electrical pulses that set up a certain state in the memory circuits (more semiconductors) of a digital computer. The computer also receives a train of signals from a second complex of electronic devices that purports to measure the temperature of the black body. Finally the computer, having itself issued electronic instructions for changing the temperature, calculates a relation between temperature and light intensity. It would be difficult to argue that the result is independent of theory (see also the similar analysis in Toraldo di Francia, 1981). Hesse's argument might seem to apply equally forcibly to the realisation of the standard of frequency which is described below in this chapter and the next.

All observation or experiment involves three elements, a purpose which arises from some prior theoretical issue, physical instruments and procedures, and an abstract model of those instruments and procedures with which the 'result' of the observation or experiment is calculated. Hesse's argument is that instruments and procedures themselves as well as the interpretation of their readings, depend on pre-existing theory, and that can hardly be gainsaid. Furthermore, as Pickering (1989) has insisted, observations as the outcome of experiments incorporate available instrumental techniques as well as the theoretical underpinning.

When the 'result' of an experiment is finally established it may or may not agree with the theoretical scheme which prompted the observation in the first place. One obvious reason for that is that the scheme is an inadequate representation of nature, that the observation encounters what Pickering (1989) calls 'resistances'. In that case the observations provide new knowledge about the natural world. The 'result' may also disagree with the initial theoretical scheme because the model used to derive it from the raw data does not correspond closely enough with the physical processes and relations. Observations are never interpreted independently of some abstract model of the physical system, analyses and calculations of results are done on the model quantities which are supposed to correspond to the physical quantities. Thus we make some calculations with a 'voltage'. We do not observe a 'voltage' directly, but rather some symbols on a digital voltmeter that are intended to correspond to the abstract notion of 'voltage' as established by the theory of the instrument and the manner of its construction.

Even in the very simplest cases, abstract representations of the physical state are involved, as in the calculation of the volume of a nominally regular solid from measurements of its dimensions (see Cook, 1961, 1975). The model used is the relation between volume, V, and the position vector, $\mathbf{r}$ of an element $\mathrm{d}S$ of the bounding surface:

$$\int_S \mathbf{r}\cdot \mathrm{d}S = \int_V \operatorname{div}\mathbf{r}\,\mathrm{d}V = 3\int_V \mathrm{d}V,$$

or

$$V = \frac{1}{3}\int_S \mathbf{r}\cdot \mathrm{d}S.$$

That simple result can be deceptive: it assumes that we know the position vector for each surface element, measured from the same origin, but since it is usual to measure distances, $\Delta\mathbf{r}$, *between* surface elements, for example across diameters of a supposed sphere, hidden shifts of origin can occur leading to erroneous calculations. Those and other possible discordances are not negligible if a volume is to be determined to one part in ten million. The discrepancies that arise as a result of an inadequate model of an experiment or observation are what we call systematic errors, and it is well known that they can be very difficult to identify.

Analyses of the reasons for making observations, of the ways in which they are made, of the theory and technology on which they depend, are no doubt of considerable interest, but they do not of themselves invalidate the results of observation or theory. Indeed it can be argued that far from casting doubt on an experimental result, the fact that a procedure is based on well established theory gives an assurance that the result is telling us something about the real natural world (see Franklin, 1989).

I assert, however, that analyses of that sort are irrelevant to the argument that I shall develop in this book. I take the results of observation as a physicist presents them, for however they were obtained they are the empirical basis of physics, or as Toraldo di

Francia (1981) puts it, *a physical quantity is defined by prescribing the operations that are carried out in order to measure it.* I seek to understand how the results of the actual practice of observation, the data as they are, determine the structure of theory. Theory in the first place must bring order into the results of observations that physicists carry out and my main purpose is to see how the one determines the other. Whether or not theory then tells us something about the real world behind the observations is a question that I defer to the final chapter.

I adopt in this book the concept of a theory as a model of our observations of the real world – not a model of that real world itself which, as I have asserted, is not directly accessible to us, but explicitly a model of the results of observations of the world, with observations defined by the operations which produce them as I have just explained. There is a view of theory, the 'instrumentalist' which maintains that a theory is just a means, an instrument, a way of calculating the outcome of observations, and that the content and structure of the theory do not necessarily bear any relation to the independent world behind the observations. An operational view of observation and an instrumentalist account of theory, while they are evidently consistent, do not necessarily entail the one the other. I do adopt throughout the operational account of observation but I consider that a true theory is more than just a calculating machine. I take a theory to be a mathematical realisation of an abstract system that has properties corresponding to those of a set of observations, a concept which I shall develop in subsequent chapters and that is at the heart of much of my argument. It is in that sense that I take a theory to be a model of the world of observations, with the implication that there is a more fundamental correspondence than just giving the right answers, and my aim in this book is to show how that fundamental correspondence comes about. That view of theory might superficially seem to be similar to Plato's notion of the relation of our world of sense impressions to an ideal mathematical world. There is, however, a deep distinction, for Plato considered the physical world that we experience to be an imperfect realisation of the more real formal structures of the mathematical world, whereas my position is that the physical world is primary and the abstract system is the best we can do to represent it.

1.3 Measurement and standards

My concern in this book is with the objective factors in the interaction between nature and the observer, of which one of the most important is the process of measurement (see Cook, 1977, 1992). Measurement is the basis of all physical science and the consequences of the constraints that it imposes are the topics of the next two chapters. I therefore go on to summarise the nature of measurement, to describe the system of standards on which physical measurements are based, and to relate them to the equations of physics.

All equations of physics, for example, Newton's equation of motion, $\partial \mathbf{v}/\partial t = \mathbf{F}/m$, are representations of physical states or processes in that they are mathematical relations that are congruent to the relations between observations. Measurements are necessary to establish the correspondences and to ensure that the representations of reputedly similar observations are compatible.

Every measurement consists of comparing some quantity with a standard quantity of the same type, and thus assigning a number to the measure of the unknown quantity in terms of the standard. Lengths are measured by setting objects alongside other objects on which standard lengths are marked out. Times and frequencies are measured by comparing them with times or frequencies of electrical signals derived from some standard oscillator. Masses are measured by comparing them on a balance with standard masses. Electrical voltages and currents are measured by setting them directly against voltages and currents supplied by standard sources. The precision of any measurement is determined both by the accuracy of the comparison with the standard and by the precision with which the standard quantity can be realised and reproduced. The system of standards of physical quantities affects all physics and all applications in engineering, and it will be argued in the next two chapters that the ways in which we measure and the choice of basic standards determine also some of the fundamental structures of physics. The nature of the system of basic standards is crucial to the argument and to it I now turn.

It is well known that there is no need to have a separate standard for every physical quantity and that in fact standards for all physical measurements can be derived from just four independent standards, conventionally those for time, length, mass and electrical current. The Système International des Unités (SI) has as its fundamental independent units the Second, the Metre, the Kilogramme and the Ampère, but that statement is already somewhat out of date, if not misleading, for it implies that there are indeed distinct physical objects, such as the standard metre, by which those standard units are realised. In fact the standard of length is derived from that of frequency by an independent value of a fundamental constant, the speed of light. The unit of length is the distance travelled by electromagnetic radiation in free space in a specified time, and its value in terms of the conventional metre is derived from the standard of time and a conventional value (2.997 924 58 m/s) adopted internationally for the speed of light (Appendix – Resolution A4 of the XXI General Assembly of the International Astronomical Union, 1991). The particular numerical value ensures that lengths derived with it are consistent with those derived from earlier physical standards, but the precision with which lengths can be derived from light times is greater than that of realisation of the now superseded physical standards of length.

Similarly the standard of electrical voltage can be derived from the standard of frequency through the Josephson effect and a conventional value of the ratio h/e of Planck's constant to the electronic charge. Here again the precision with which a voltage can be derived in that way is better than the accuracy of the value of the ratio h/e in the terms of the electrodynamical standards of electrical units. The standard of mass remains as yet unrelated to the standard of frequency but since it is now possible to relate the unit of electrical current to that of voltage through the resistance of the quantum Hall effect, it is conceivable that the unit of mass could be replaced by a unit of energy derived from the units of electrical current and voltage and hence related to the standard of frequency through the two constants of the Josephson effect and the quantum Hall effect.

Nowadays then, all the other fundamental units can be related to the standard of frequency through conventional values of certain constants of physics. It should be appreciated that relating other quantities to frequency has not reduced the number of independent quantities on which the system of measurement is based. Although we no longer use an independent metal bar to realise the standard of length, the conventional value of the velocity of light that we use is equally arbitrary and is as much an independent physical element of the system of standards as the metal bar.

The standard of frequency is itself realised by an atomic process and two of the three constants arise from quantum processes, the Josephson effect and the quantum Hall effect. Quantum physics has thus to a large extent replaced classical physics as the basis of the standards of measurement (Petley, 1985). The reason for that is nothing very subtle, it is simply that the resulting system of standards, units and measurement, depending as it does on various electrical and electronic measurements, is more convenient in use, more precise and more generally accessible, than the mechanical system it has replaced. At the same time, the implications for the logical structure of physics are profound, as will be argued in the next two chapters; or rather, the use of the new scheme of units and standards reveals that structure more clearly than may often have appeared in the past.

1.4 The standard of frequency and standards derived from it

Before drawing out the implications of the new system, the physical nature of the standard of frequency and of the constants of physics, as well as their logical position in the scheme of physics, must be explained. The present standard of frequency is the frequency of an electrical signal that causes transitions between the two hyperfine levels in the ground state of the atom caesium-133 (Appendix – Resolution I of the XIII CGPM, 1967). The standard is realised physically in apparatus in which the atoms in an atomic beam of caesium-133 first pass through a magnetic filter that prepares them in their hyperfine states as distinguished by their magnetisation. They then pass through a region in which an electromagnetic field is maintained at the correct frequency (9192.6 MHz) followed by a second magnetic filter to detect when transitions between the hyperfine levels have occurred. It is found in practice that the frequencies of the electrical signals realised in that way in different laboratories agree to within about 1 part in

10^{13} or better. The standard is therefore highly reproducible; it is also widely and easily accessible through radio transmissions. A somewhat more convenient apparatus is the atomic hydrogen maser, in which an inversion of the populations of the two hyperfine levels in the ground state of atomic hydrogen is brought about by a magnetic filter. Stimulated emission from the upper level maintains electrical oscillations in a cavity tuned to the microwave frequency of the transition, about 1420 MHz. The maser is convenient because it generates a continuous electrical signal, but the frequency depends to some extent on coupling with the microwave resonant cavity and so is considered to be less fundamental than that of the caesium standard (Kartaschoff, 1978).

It is evident that both forms of standard depend heavily for their design and operation on theory, not only for the basic principle, but in the operations of the source, the detector and the filter. None of that invalidates the status of the apparatus as a means of realising a fundamental standard, for by an internationally agreed definition, the standard frequency is the frequency generated or identified in the operation of that apparatus. All that is necessary is that design, construction and operation of the apparatus should be so closely specified that everyone who operates an example of it to the specification should get consistent results.

Once a standard has been defined for some quantity, it is meaningless to speak of how it may change or of checking it against some other standard. The second was originally defined by the rate of rotation of the Earth upon its axis and when it was suspected that the Earth's rate of rotation might vary, the second (now the ephemeris second) was re-defined in terms of the period of the Earth in its orbit about the Sun. It then became meaningful to talk about the variable spin of the Earth whereas previously it had not because there had been no better standard against which to test the spin. We have to recognise, as Wittgenstein has emphasised, that some apparently well formulated questions have no answers. Now that the second is defined by an atomic process, we can in turn discuss the possible variation of the ephemeris second, but it is meaningless to speak of changes in the atomic standard itself unless an improved way of defining the second is developed and replaces the caesium standard by general consent. There is of course a practical question of how to define the standard frequency when