OXFORD

R. J. Hankinson

CAUSE AND EXPLANATION IN ANCIENT GREEK THOUGHT

Cause and Explanation in Ancient Greek Thought

Hankinson, R. J, University of Texas, Austin

Abstract: In this book, R. J. Hankinson traces the history of investigation into the nature of cause and explanation, from the beginnings of Ancient Greek philosophy in 600 bc, through the Graeco-Roman world, to the end of pagan antiquity in c.500 ad The book consists of chapter-length studies of the Presocratics, Plato, Aristotle (two chapters), Atomism, Stoicism, Scepticism, and Neoplatonism, as well as the Sophistic movement, and Ancient Medicine. Hankinson is principally concerned with the following questions: 'What did the Greeks understand by a cause?', and 'How did the Greeks conceive adequacy in explanation?'. The Ancient Greeks (excepting the Sceptics) are united in their belief that the world and at least some of its process can be rendered intelligible, and that this can be rendered by an inquiry into the nature of things, with reasoned argument as the appropriate method of exhibiting the real structure of the world. Thus, the Greek thinkers set the standards for science, because they are guided by logic and observation in their analysis of causation; but one can also recognize the growth of interest among the Greeks in the nature of explanation itself. The question that becomes central to the development of Greek philosophical science is whether nature can be understood in terms of teleology, or solely in terms of mechanical laws. Hankinson is interested in how the concepts of cause and explanation function in a properly scientific context; but he extends his investigation of these concepts to questions of freedom and responsibility, and fate and astrology, and also the treatment of disease. Hankinson points out that causes and explanations are connected ideas: an explanation is the proffering of reasons, and this involves an account of causes; they are, nevertheless, different concepts—causes are actual items, events, agents, facts, states of affairs, whereas explanations are propositional. Hankinson isolates certain causal principles that recur throughout Greek philosophy: for instance, the principle of sufficient reason, the principle of causal synonymy, and the principle that nothing can come to be from nothing.

Preface

It was first suggested that I write this book in September 1986: it has, accordingly, taken more than a decade to come to fruition. I began by supposing that it would consist in a relatively narrowly focused investigation of the Greek notion of explanation; it has developed into something more like a general account of what the Greeks called *phusikē*, the inquiry into nature, since I found it impossible to treat of the meta-theoretical issues of how the Greeks viewed the concepts of explanation and cause themselves without also considering the actual explanations and causes that they championed (it also accounts for my ranging beyond the boundaries of natural science and into metaphysics). This partly accounts for its elephantine period of gestation; but I did not in any case set about the writing in earnest until the spring of 1992, when I produced a first draft as work in progress for a graduate seminar: I am grateful to the seminar's students for their patience and flexibility in adapting to an unorthodox format, and for their enthusiasm and input.

My interest in questions of cause and explanation dates from the time when, as an undergraduate, I decided to pursue these issues in Galen's philosophy for my doctoral dissertation, and they have continued to be central to my work in Greek philosophy and science. Inevitably, then, much of this volume goes over ground I have covered elsewhere, although nowhere have I simply recycled previously published material, while in several cases my views have changed. Still, it may be appropriate to indicate areas of overlap.

I dealt with Alcmaeon's argument (I.3*c*) in Hankinson (1991*a*, 1992*a*), and some of my treatment of Xenophanes in I.3*a* is paralleled in Hankinson (1995*a*, ch. iii), as is my account of Democritean scepticism in VI.1*e*. Some of the Hippocratic discussion of Chapter II is anticipated in Hankinson (1991*a*, 1992*c*, 1995*e*), while the remarks on evidence are expanded in Hankinson (1997). Plato's Alcmaeonian argument in III.3*a* is developed in Hankinson (1992*a*). Much of the discussion of Aristotle in Chapters IV and V is adapted from Hankinson (1995*b*, *c*), while the Atomist and Stoic accounts of causation and explanation that occupy Chapters VI and VII are rehearsed in Hankinson (1998*a*, *b*); also relevant is (1996). The sceptical attack on explanation and cause, the subject of Chapter VIII, is also dealt with in Hankinson (1995*a*, chs. xi and xii); while ch. xiii of that work, along with Hankinson (1987*b*, 1988*d*, 1990*a*, 1995*d*, 1998*d*), as well as parts of (1998*c*), all bear on the discussion of the medical tradition in Chapter IX. Hankinson (1988*a*, *c*, 1995*a*, ch. xv) relate to the

issue of the status of divination discussed in VI.3*a*, VII.2*d*, VIII.5, XI.1*b*, and XII.2*d*. Finally, my account of Galen's views in XI.2–5 draws on Hankinson (1987*a*, 1988*b*, *d*, 1989, 1991*a*, *b*, *c*, 1993, 1994*a*, *b*, 1998*c*). Many of these articles were originally presented as papers in various places; I owe much to the perspicacity of their original audiences.

Particular thanks are due to several colleagues and friends for their help with individual sections. Alex Mourelatos gave me the benefit of his great learning in Presocratic philosophy in general and atomism in particular (Chapters I and VI); Lesley Dean-Jones commented acutely on the medical material contained in Chapters II and IX; and Paul Woodruff helped me with both Plato and the Sceptics (Chapters III and VIII). Sylvia Berryman made several useful suggestions concerning Theophrastus and Strato (V.3). The argument of Chapter XII owes much to discussions with David Bradshaw, who was also good enough to read and comment in detail on penultimate drafts of that chapter and Chapter X. My largest debt in this regard is owed to Tim O'Keefe, whose doctoral work on Epicurus and Democritus on reduction and emergence, causation and responsibility, determinism and fatalism, caused me to re-evaluate and revise many of my views on these issues, particularly as they affect the argument of VI.1*b*, 2*b,* and 3. It is both a pleasure and a privilege—and one which it is all too easy to take for granted—to work with such generous and genial colleagues.

I am grateful also to Peter Momtchiloff of Oxford University Press for taking the book over after the series in which it was originally due to appear was peremptorily cancelled, without either consultation or explanation, by another publisher. An anonymous referee

for OUP made several valuable suggestions which I have incorporated into the final version.
Finally, my greatest indebtednesses. In the course of a long, presumably insomniac, European rail journey, Jonathan Barnes read and commented upon the whole typescript with admirable expeditiousness and acuity, suggesting numerous improvements ranging from the correction of typographical errors through matters of fact and questions of argumentative substance and rigour to issues of overall strategy and presentation. His influence is evident (to me at least) on virtually every page.
And last, but most definitely not least, my wife, Jennifer, read the manuscript in full, offering many suggestions of both a philosophical and of a stylistic nature. I am also grateful for her support (not to say understanding and forbearance) through a period in which both of us have had to juggle heavy workloads with family responsibilities. Anyone who has had to attempt such feats of domestic dexterity will know what I mean when I say that merely to survive is an achievement; to do so with good will
end p.viii

and good humour intact is little short of miraculous. In recognition of this (and much else), with love and gratitude, I dedicate this book to her.

Jim Hankinson
Texas, April 1997

Note on Citations

I have tried to make my citations of ancient texts as full and perspicuous as possible. Where a referential orthodoxy exists, I have followed it; in default of one, I have sought to make it as easy as possible for the interested reader to track down the sources.
A number followed by the initials KRS refers to that fragment in G. S. Kirk, J. E. Raven, and M. Schofield, *The Presocratic Philosophers* (Cambridge, 1983). Similarly, DK refers to H. Diels and W. Kranz, *Die Fragmente der Vorsokratiker*, 6th edn. (Berlin, 1951). DK references consist of an initial number, picking out the philosopher in question (thus Thales' citations all begin within the number 11), followed by a letter, A for indirect reports and testimony, B for actual fragments. I give DK references to the A-texts only in the case of the more important citations. 'SVF' flags H. von Arnim, *Stoicorum Veterum Fragmenta*, 3 vols. (Leipzig, 1903–54). 'Jones' after a Hippocratic text indicates that the citation refers to W. H. S. Jones, *Hippocrates*, i, ii, iv, Loeb Classical Library (London, 1922, 1923, 1932). 'FHS&G' refers to W. W. Fortenbaugh, P. M. Huby, R. W. Sharples, and D. Gutas, *Theophrastus of Eresus: Sources for his Life, Writings, Thought, and Influence* (Leiden, 1992). 'W' refers to F. Wehrli, *Straton von Lampsakos*, 2nd edn. (Basel, 1969). 'Walzer' refers to R. Walzer, *Galen on Medical Experience* (Oxford, 1944). 'LS' refers to A. A. Long and D. N. Sedley, *The Hellenistic Philosophers*, 2 vols. (Cambridge, 1987). 'Bruns' refers to I. Bruns, *Supplementum Aristotelicum*, 2. 1–2 (Berlin, 1887, 1892), which contains Alexander of Aphrodisias' peripatetic essays. 'U'

refers to H. Usener, *Epicurea* (Leipzig, 1887). ‘EK’ refers to L. Edelstein and I. G. Kidd, *Posidonius*, i: *The Fragments* (Cambridge, 1972). References to Galen are generally given by way of the volume and page numbers of the edition of C. G. Kühn (1821–33), even where later and better texts exist (since these later texts also contain the Kühn pagination); for texts not edited by Kühn, other references are supplied at the appropriate junctures. ‘vS’ references are to the collection of testimonia and fragments of Herophilus in von Staden (1989). ‘G’ refers to I. Garofalo, *Erasistrato* (Pisa, 1988), a partial collection of Erasistratus' fragments. ‘H’ refers to G. Helmreich, *Galeni Pergameni Opera Minora*, iii (1893, Leipzig); ‘D’ to K. Deichgräber, *Die griechische Empirikerschule. Sammlung und Darstellung der Lehre* (Berlin, 1930); ‘Walzer’ to R. Walzer, *Galen on Medical Experience* (Oxford, 1944). ‘IP’ refers in the case of Speusippus to M. Isnardi Parente, *Speusippo: Frammenti* (Naples, 1980), and in the case of Xenocrates to M. Isnardi Parente, *Senocrate,*
end p.xv

Ermodoro: Frammenti (Naples, 1982). ‘Festa’ refers to N. Festa, *Iamblichus: de Communi Mathematica Scientia Liber* (Leipzig, 1891). Fragments of Numenius are cited by way of E. des Places, *Numenius: Fragments* (Paris, 1973). ‘Dodds’ refers to E. R. Dodds, *Proclus: The Elements of Theology* (Oxford, 1963); ‘Wildberg’ to the fragments of Philoponus'*Against Aristotle on the Eternity of the World*, collected in C. Wildberg, *Philoponus: Against Aristotle on the Eternity of the World* (London, 1987).

Additional Note

In the text I make considerable use of abbreviations to designate formally expressed principles. Where they occur subsequently in the text, I have tried to refer back to their first occurrence and explanation; but for convenience, I have collected them, in the order in which they occur, along with brief explanations of their function, in the Appendix (‘List of Abbreviated Principles’), to which the frustrated reader is referred.
end p.xvi

Introduction

R. J. Hankinson

Why do things happen? What makes an event occur at a particular time? What are the ultimate constituents of things? How does the structure of some organism account for or determine its function? Are the constituents of the universe there for some purpose? What is meant by chance and coincidence? Are there such things as natural laws? How are people responsible for what they do?

Although these questions vary in sophistication and scope, they are united by the fact that they are demands for explanation. They are questions that call for an elucidation of the structure of the world, for an account that will render its apparently arcane and random processes amenable to prediction and control. They are the fundamental questions in the sciences of physics, chemistry, and biology, as well as in metaphysics and ethics.

There is, however, another feature which they have in common: they were all posed, in a variety of different forms, and often with widely varying motives, by the Greeks (see also Hankinson 1988*c*). Perhaps the Greeks did not invent causal explanation; but they were certainly the first Western civilization to subject the ideas of cause and explanation to rigorous and detailed analysis, and to attempt to construct, on the basis of both logic and experience, grand theories about the relations that hold among the inhabitants of the physical universe. That is to say, they engaged in both science and the philosophy of science (on the interplay between reason and experience, see Geoffrey Lloyd's magisterial book *Magic, Reason and Experience*; G. E. R. Lloyd 1979).

This book is an attempt to trace the Greek history of these ideas from their earliest beginnings, through Plato and Aristotle to the Hellenistic philosophers, arriving finally at the Neoplatonists, more or less (the treatment of the Atomists being the principal exception) in chronological order. This is a large and somewhat unwieldy project, which has (in exemplification of a causal principle dear to the Greek soul) resulted in a large and somewhat unwieldy book. My principal concern has been to pursue the development of the concepts themselves: what sort of thing did the Greeks think causes were, and how did they conceive of adequacy in explanation?

Those questions suggest that the Greek theorists themselves were

actively conscious of the intricacies of the ideas involved; but such an awareness is not really evident before Plato and Aristotle. None the less, earlier thinkers implicitly presuppose that causation and explanation must have a particular structure, and it has proved impossible to tackle the main theme of the book without spending a considerable amount of time in examining how the Greeks actually did seek to explain things, prior to an analysis of their self-conscious philosophical reflections on the notions. I begin in earnest with the Presocratics (Chapter I), since I share Aristotle's view that they instituted a revolution in the way the world and its workings were conceived, and were indeed, if not the first scientists, at least the first theorists whose attitudes exemplified the detachment as well as the speculative boldness of the scientific temper.

It is a matter of dispute, much of it sterile and pointless, whether we should attach the labels ‘science’ or ‘philosophy’ to the early Presocratics. My reasons for enlisting them under the banner of science emerge in the succeeding paragraphs. Mansfeld (1985*a*) has recently argued that gains in clarity of analysis and problem-posing are to be derived from treating the Milesians as the sires of science, and Parmenides and Heraclitus as the fathers of philosophy, on the grounds that in its modern sense at least philosophy consists in the reflection upon problems raised by the pursuit of other forms of intellectual endeavour, notably science. I am inclined to agree (weakly) with this view, with the proviso that Xenophanes too should be considered a philosopher on this score.

An account is scientific in this broad sense if it seeks to explain events by appealing to repeatable and generalizable laws, laws that are invariant over time (or at least suitably long stretches of time), and which can be used to ground predictions of future occurrences. Such views are naturalistic: events are explained as the natural consequences of particular antecedent circumstances. By contrast, pre-scientific thought places less emphasis on repeatability and predictability, and tends to appeal to supernatural forces, divine interventions, the malevolence of demons, and other such occult apparatus to account for the way things are and the sorts of event that occur.

Of course, in a perfectly straightforward sense, such pre-scientific world- pictures deliver explanations of sorts. If you want to know why there has been an earthquake, a seer, no less than a seismologist, can give you an answer to that question. Their answers will, however, differ markedly in type. The seer will ascribe the disaster to divine displeasure at human injustice, blasphemy, or impiety; the seismologist will talk of plate tectonics and movement along fault-lines. The issue, then, is not whether or not to offer explanations at all, but rather what kinds of explanation can and should be essayed.

The basic features which distinguish scientific from other sorts of

explanation are their universality and simplicity. Science is concerned not with the explanation of some individual event, or some particular ephemeral occurrence, but rather with producing a quite general account of why things *of this sort* happen the way they do *in this type of circumstance*. We do not offer bespoke explanations for individual earthquakes (Poseidon was angry here; Zeus displeased there). If our scientific explanation of a particular earthquake is to be scientifically adequate, it must in principle be capable of explaining why any earthquake occurs. This is, as it stands, too strong a condition: a science of seismology may be genuine and predictive even if it can only account for a limited subclass of all seismic activity. But it must be able to deal exhaustively with that class, as well as explaining why the subclass in question is a proper (i.e. explanatorily unified) subclass; and these are not trivial demands. Another (perhaps tendentious) way of putting that is to say that scientific explanation works by subsuming individual events under general laws. What the science consists in is just an exhaustive enumeration of the laws and regularities that are instantiated in that particular area of inquiry.

Laws and regularities need not take the form of the contemporary mathematical conception of a scientific law, one in which (whatever your metaphysical predilections) physical quantities are related by general functions that establish their co-variability, along the lines of the classical modern account of the relations between law and explanation offered by Hempel (1965) (of course, that is far from being the only contemporary account of scientific explanation; for a useful contemporary *dégustation*, see Pitt 1988; Ruben 1990). Something of that kind of course will certainly count as a scientific law; but so, on this account, will something looser and less rigorous, namely any assertion of a universal, or even general, truth: 'most men have beards', to borrow an Aristotelian example, falls into this category.

Simplicity is, paradoxically, a notoriously difficult concept. Broadly, an explanation *E* of some set of events *S* is simple just to the extent that *E* can account for the greatest possible variety of members of *S* on the basis of the fewest assumptions. Simplicity, then, is an adjunct of generality: the greater the generality of an explanation, the greater its simplicity. Simplicity and generality are connected with reductivism. An explanation is reductive just in case it accounts for a diversity of phenomena on the basis of a limited number of fundamental postulates; it is also reductive to the extent that the phenomena are said to be caused by, or based on, those postulates, which in turn explain the phenomena.

Explanation standardly takes the form of stating the causes of something; and a scientific explanation is in principle no different in this respect from any other explanation. What is distinct is the type of cause adverted to. To state the causes of something in a scientific way is to point to the

nature of things: those basic truths about the world in terms of which our account of its workings is to be structured. A scientific explanation, then, will pick out those more fundamental facts about the structure of the universe on the basis of which the things to be explained can be explained. Science is, then, inherently reductionist: it is the enterprise of seeking to account for the greatest variety and multiplicity of phenomena on the basis of the fewest fundamental postulates: and the Presocratics grasped, if tenuously, that basic fact.

Finally, scientific explanation involves reason. What distinguishes the early Presocratic physicists from their predecessors is, among other things, their use of argument to support their theoretical claims. And while they may seem to us absurdly optimistic about their arguments' probative force, and hopelessly naïve in their deductive techniques, what matters is that they employ them at all. And by Aristotle's time arguments (in the sense of deductive chains of reasoning) have come to be the appropriate method of exhibiting the real structure of the world. The process by which, in a completed science, you deduce the theorems about the subject-matter from prior and more basic axioms mirrors (or at least should do so) the actual hierarchical structure of reality.

I should at this point enter a caveat: all my discussion of science has been couched so far in severely and perhaps tendentiously realist terms; progress in science is a matter of providing ever more accurate pictures of reality. That is, of course, not the only way in which science can be viewed (although it was the one overwhelmingly adopted by the ancient physicists); and Greek mathematical astronomy in particular (which I deal with briefly in V.2*f*, X.2*b*, XI.1*a*) demonstrates that the ancients could, on occasion, treat scientific theories merely instrumentally, as predictive models. But such views were exceptional (see G. E. R. Lloyd 1978*a*).

I have deliberately vacillated between talking of causes and of explanations. The two ideas are clearly connected: when we explain something, we generally give reasons for its being the way it is; and to give reasons why something is the way it is frequently involves an account of the causes of that thing. Conversely, by specifying the causes of some event or state of affairs, we are inclined to think that we have gone at least some of the way towards explaining it. But for all that, the two concepts are different, even if closely related: and a good deal turns exegetically on whether one chooses to render the Greek word *aition* (or *aitia*) as 'cause', 'reason', or 'explanation'. Causes are actual items, events, agents, facts, states of affairs; explanations, on the other hand, are propositional. Furthermore (and relatedly) causal contexts are extensional, while explanatory ones are not (this latter claim is controversial, and is rejected by Anscombe 1969; but see Mackie 1974, ch. 10; see also Sorabji 1980*a*, chs. 1–3).

Moreover, it is important to distinguish between an analysis of causal language as it is actually employed, and self-conscious theories of cause and explanation. None the less, when we reflect upon the variety and proliferation of causal language, on the range of cases and types it is supposed to cover, the need to establish some analytic order among the chaos rapidly becomes apparent; and, I shall be arguing, this need was no less obvious to the ancient Greeks than it should be to us. So a concern with providing explanations quickly generates a second-order concern with the nature of explanation itself; and that concern is evident beneath the surface even when such methodological questions are not, as they are not prior to Plato, directly broached.

For all the (very real) differences between the scientific and the mythographic temper, it is no part of the former to disavow religion altogether. From Aristotle and Galen through Kepler and Newton to Einstein and Eddington and beyond, there have been plenty of scientists of a religious, not to say mystical, bent. What differentiates the religious scientist from the mythographer is not belief or unbelief, but rather a disagreement about the scope of purely religious explanation. Newton elaborated laws of very great generality to account for the observed motions of the heavenly bodies (and a variety of other phenomena): but he did not, so he said, frame hypotheses, by which he probably meant that his laws were intended only to describe the functioning of masses *as they are*: they were not designed to explain *why* things *should* (in some strong sense) function the way they did, nor to account for how the initial dynamic impetus was supplied to the cosmic machine: those were questions for theology. Newton was following Galileo's lead: the job of science is to describe how, not to explain why. (This does not of course debar scientists from trying to construct arguments in natural theology, arguments designed to show that scientific facts somehow provide evidence for the existence of some sort of supreme being, as they have done periodically throughout the history of science—indeed there is apparently a current revival of interest in such enterprises.)

But that was by no means apparent to the Greeks; and throughout their philosophy of science we may discern a preoccupation with the question how much, if at all, it is appropriate for the rational investigator of the universe's inhabitants and processes to seek to explain them in terms of the ends, goals, or purposes they subserve. Is the universe a complete structure, unified by the way in which the various functions of its parts are conducive to its overall operation (is it, then, as Plato and the Stoics believed, a kind of animal, and under the direct control of God)? Or is it merely the functional outcome of an agglomeration of initially random and undirected mechanical particular interactions (as the Atomists were to argue)? Or is there, as Aristotle held, some coherent middle way between the view that the cosmos is the work of God on the one hand and the purely mechanistic physics of chance on the other? Can Nature itself, intrinsically, exhibit a type of purposiveness which serves to explain its regularity and stability without appealing to some super-intelligence?

The thinkers celebrated in this volume are (with the exception of the Sceptics) united by their basic belief that the world and some at least of its processes can be rendered intelligible to human investigators, and that its intelligibility is to be discovered by an inquiry into the nature of things. But whether that nature is such as to be completely describable without remainder in terms of purely mechanical laws of working, or whether rather nature demands to be understood in teleological terms, is the central question of Greek philosophical science, and one which affords, albeit sinuously and sometimes obscurely, this inquiry's unifying thread.

I The Presocratics

R. J. Hankinson

1. Prehistory

The earliest surviving literary products of Greek civilization are the Homeric poems. These, the outcome of a long oral tradition, were probably first written down in the eighth century bc ; but they had been in gestation for perhaps two hundred years. They contain passages in which primitive theories of the structure of the universe can be discerned. But they are not the first such notions to appear in the history of civilization, and anthropological investigations confirm that virtually all societies have their creation myths. The Babylonians generated the world and its contents from Apsu, fresh water, and Mummu-Tiamat, the sea and cloud;[1] in the principal Egyptian cosmogony, the primordial waters (Nun) bring forth Khephri, the sun-god, who sees to the generation of everything else (Pritchard 1969: 6).[2]

These cosmogonies date from the third millennium bc , and are thus greatly older than the Homeric poems. But the latter too make water in a way basic; the earth is surrounded by the great Stream of Ocean, from and into which the heavenly bodies rise and set (*Iliad* 18. 607–8, 21 194–7 = 4, 6 KRS; cf. 7. 422).[3] Some (admittedly isolated) passages go further, and suggest in the manner of ancient Near Eastern cosmogonies that Ocean was the begetter of the gods and the rest of the world (*Iliad* 14. 200–1, 244–8 = 8, 9 KRS). Both Plato (*Theaetetus* 152e) and Aristotle (*Metaphysics* 1. 3. $983^{b}27$–$984^{a}1$) read serious metaphysical significance into these passages, the former claiming that it showed Homer to believe that ‘everything was the offspring of flux and movement’. The sky was conceived of as a hollow brazen hemisphere (*Iliad* 17. 425). The atmosphere is divided into two regions, the lower being the moist, misty *aēr*, the upper the clear, shining *aithēr*, the abode of the gods. The gods, in ultimate if sometimes quarrelsome subservience to Zeus, direct human affairs according to their whims and prejudices, but in the final analysis in response to the will of Zeus, the ineluctable working-out of *moira*, fate or destiny.

It has sometimes been suggested that the divine mechanism of the Homeric poems is simply intended as metaphorical ornamentation, and as such should not be taken seriously; but there is little reason to believe this.[4] Earthquakes are caused because Poseidon, 'the earth-shaker', is displeased with some segment of humanity. Lightning and thunder are the arms of Zeus. The appeal to massive supernatural forces in order to account for large-scale physical events is explanation of a sort: but it appears piecemeal and *ad hoc*, and does not lend itself to general formulation in the form of useful physical laws,[5] even had the Archaic Greeks thought in such terms, as they did not. Equally, the cosmogonies and the structural accounts cannot be subjected to empirical test, and no evidence is adduced in favour of them.
Of greater importance than the scattered passages in Homer, or the later references to early 'Orphic' cosmogonies in some of which water also figures,[6] is the poem by the seventh-century Boeotian poet Hesiod, the *Theogony*, or birth of the gods. In it, the poet recounts how first of all there was Chaos, a gap or chasm—and from it came the earth and Tartarus, then Eros, 'fairest among the immortal gods, the limb-loosener'; and then night and day, the air and sky, the ocean, and finally the physical features of the earth itself and the gods (*Theogony* 116–38). This is not science by any stretch of the imagination; and it comes from the same intellectual stable as the Near Eastern cosmogonies. But it evinces a concern for the ultimate explanation of things which was to characterize Greek thought.

2. The Milesians

(a) Beginnings

The Greeks themselves (Aristotle, *Metaphysics* 1. 3. 983^{b}20 = 11 A12 DK) were inclined to think that philosophy began with Thales of Miletus, who was probably born in the last quarter of the seventh century. His contemporary Pherecydes of Syrus is sometimes mentioned alongside him; Aristotle found in him an amalgam of myth and serious natural science, although precious little of the latter survives. His cosmology featured as eternal, ungenerated principles 'Zas [i.e. Zeus], Time, and Chthonie [the earth]' (7 B 1 DK = 49 KRS); Aristotle refers to his 'non-mythological' talk of a Supreme Good (*Metaphysics* 14. 4. 1091^{b}8–11).
end p.8

Of Thales we know virtually nothing, most of what we find in later sources being the fictitious constructions of a degenerate biographical tradition. Hence, *pace* Diogenes Laertius (*Lives of the Philosophers* 1. 23 = 11 A1 DK: reporting Eudemus of Rhodes), he is unlikely to have discovered general geometrical theorems. He may have had practical geometrical skill: he is said to have calculated the height of the pyramids by primitive triangulation (ibid. 1. 27 = 11 A1 DK), and to have been a consulting engineer for Croesus' Lydian army (Herodotus 1. 75 = 11 A 6 DK). Most famously of all, he was supposed to have predicted the eclipse of the sun of 585 bc (id. 1. 74 = 11 A 5 DK), although he cannot have done so with any great precision (if indeed he did at all). Thales was no master-astronomer or mathematician.

(b) Material Monism?

But however apocryphal these tales, the Greeks themselves saw Thales as a towering figure in their intellectual history. Aristotle makes Thales the 'founder of a certain type of philosophy' (*Metaphysics* 1. 3. 983^{b}20–1 = 11 A 12 DK = 85 KRS), namely that which attempts to isolate the ultimate constituent or constituents of the material world. Indeed, Aristotle saw that drive as characteristic of his predecessors, and in a programmatic passage at the beginning of his *Physics*, he offers a serviceable, if (as we shall see) somewhat tendentious, taxonomy of such an enterprise:

1 there must be either (A) one principle (*archē*) or (B) many; and if (A) one, it must be either (i) unchanging (as Parmenides and Melissus say) or (ii) changeable (as the natural scientists say, some declaring (*a*) air, and others (*b*) water, to be the first principle). But if (B) there are many, they must be either (i) finite or (ii) infinite in number; and if finite but more than one, then either (*a*) two or (*b*) three or (*c*) four or (*d*) some other number; and if (ii) they are infinite, they must be either (*a*), as Democritus says, different in form although the same in generic substance, or (*b*) the opposite. (Aristotle, *Physics* 1. 2. 184^{b}15–21)

In Aristotle's view, at any rate, Thales occupies slot (Aii *b*) of that categorization:

2(i) most of the earliest philosophers thought that the principles (*archai*) of all things were merely material in form. For they say that the element (*stoicheion*) and principle of things is that from which they all are and from which they first are generated and into which they are finally destroyed, its substance (*ousia*) persisting while its properties are altered. . . . (ii) Hence they believe that nothing is either generated or destroyed, since this nature is always preserved. . . . since there is some one entity (or more than one) which always exists and from which all other things are generated. (iii) However, as to the number and form of the principle they do not say the same thing. (iv) Thales, the initiator of this type of philosophy, says it is water (this is why he says that the earth rests on water), perhaps deriving this assumption from seeing that the nourishment of

everything is moist and that heat itself is generated from moisture and depends upon it for its existence (and that from which something is generated is always its origin [*archē*]). (v) For this reason, then, he acquired this belief; and also from the fact that the seeds of everything are moist, and water is the origin of the nature of moist things. (*Metaphysics* 1. 3. 983^{b}7–27 = 11 A 12 DK = 85 KRS)

Aristotle's analysis involves his own distinctive language of matter and form (see IV.1*a, c*), and is made to serve his own particular purposes, in showing the early Presocratics as being in some ways on the right track regarding explanation, even if their treatments were incomplete and jejune. Moreover, his interpretation is avowedly speculative: **2**(iv).[7] Elsewhere he writes that

3 some say that it [i.e. the earth] rests on water . . . and they say it was advanced by Thales of Miletus who thought that the earth rests because it can float like a log or something else of such a kind (for none of these things can rest on air, but they can rest on water) as if the same must not hold of the water as holds of the earth itself. (*On the Heavens* 2. 13. 294^{a}28–34 = 11 A 14 DK = 84 KRS)

Here he ascribes, albeit cautiously, two distinct claims to Thales:

(1) everything comes to be from water;

and

(2) the earth rests on water.

Let us briefly consider the less philosophically interesting proposition (2). It is advanced in order to answer a question of pressing concern to Greek cosmologists: how come the earth is stationary, and what holds it up? Aristotle is dismissive; to say that the earth is held up by water simply prompts the further question of what supports the water: (2) is no genuine solution to the question it seeks to answer.[8] But Aristotle intriguingly implies that it was prompted by the empirical observation that certain massive, heavy things float; and he suggests empirical reasons for Thales' physical and metaphysical speculations in **2**. If Aristotle is right, this is a great innovation: nowhere in the earlier cosmogonies is there any indication of an attempt to justify the basic hypotheses by inference from observable phenomena.

end p.10

These observations are not, of course, compelling evidence for Thales' views. Even if he may evade Aristotle's first objection (by supposing the underlying water to be indefinite in extension: cf. Xenophanes, § *j* below), the evidence that he did so is weak.[9] Furthermore, his explanation of the earth's rest is empirically inadequate: Aristotle remarks that the whole earth ought to behave in the same way as individual bits of it; but plainly lumps of earth sink in water. Yet (2) remains an honest attempt to answer an apparently genuine question by extrapolation from the observed properties of things. Thales may also have appealed to (2) in order to explain earthquakes (Seneca, *Natural Questions* 3. 14): this suggests that Thales was already concerned with what will become a major preoccupation of Presocratic *physics*, namely the explanation of large-scale natural events in natural terms (and which justifies Aristotle's enrolment of him in **1** among the *phusikoi*, or natural scientists).

Aristotle suggests that (1) and (2) are systematically connected ('this is why he says the earth rests on water'); but the connection is obscure. Perhaps Thales argued that earth, as an elaboration of water, must be in some way continuous with it, which would make him a material monist of the sort Aristotle describes in **2**(i): water really is the physical substrate of everything, all things being merely modifications of water. If that is right, then everything is constituted of water; at bottom everything really *is* water.

On the other hand, many scholars[10] incline to the view that Aristotle reads his own metaphysical preoccupations into Thales' altogether more naïve and metaphysically innocent speculations.[11] Perhaps he only meant to say that

(3) water is necessary for growth,

and so for generation; and hence to infer that water must be basic for the generation of the world itself. (3) is susceptible of various interpretations: perhaps earth is precipitated from the original water as a sediment; more probably Thales considered water to be the basic source of nourishment, geological creation being no different in type to biological growth. He saw no sharp differentiation between the biosphere and the rest of the universe:

end p.11

4 some say it [i.e. the soul] is mixed in among the whole universe, for which reason Thales perhaps thought that everything is full of gods. (Aristotle, *On the Soul* 1. 5. 411^a7–8. = 11 A 22 DK = 91 KRS)

Moreover, Aristotle's conjectured reasons for (1) seem rather (against Aristotle's own apparent interpretation) to support (3), and hence suggest that Thales was not a monist at all (**2**(iv–v)). Yet even if Thales was not a metaphysical revolutionary who inaugurated reductionism in the physical sciences, his conjectures about the structure and growth of the world were at least supported by generalizations and analogies drawn from empirical evidence. No matter that the conjectures outrun their evidential base, and that the analogies are at best weak inductions, at worst fantastical extrapolations. Here are the first glimmerings of the empirical method in scientific explanation. Aristotle was right to see Thales as an innovator.

(c) The Soul

Thales was evidently a panpsychist of sorts (**4**); but he is better known for another psychological doctrine:

5 Thales, judging by what they say, seems to have supposed the soul (*psuchē*) to be something motive, given that he said that the stone [i.e. magnet] had a soul since it moved iron. (Aristotle, *On the Soul* 1. 2. 405^a19–21. = 11 A 22 DK = 89 KRS; cf. Diogenes, *Lives* 1. 24 = 90 KRS)

Thales held

(1) Magnets possess souls;

and he did so on the basis of an argument, whose premisses we may tentatively reconstruct as follows:[12]

(2) magnets have the power to cause motion;

and

(3) anything which has the power to cause motion has a soul.

(1) then follows from (2) and (3). (2) is an uncontroversial empirical fact, while (3) is, supposedly, a conceptual truth. In order to understand Thales' reasons for accepting (3) as a truth of definition, we need to examine the Greek concept of the soul, or *psuchē*.

A *psuchē* is, for the Greeks, whatever it is in virtue of which something is alive. The term *psuchē* connotes neither immortality nor dualism; and it is a trivial truth that anything which is alive has one ('soul', although conventional, is thus an infelicitous translation). The world divides into the

end p.12

living and the inanimate—and if these are genuine metaphysical categories, there must be some property or set of properties the possession of which demarcates the quick from the dead. To determine whether (3) is true, we must discover whether motion-causing is one of those properties. And it is indeed notable that one thing which does set animals off from the rest of the world is their apparent capacity to initiate motions (note that (3) does not say that everything ensouled must be a motivator: and it was later disputed, between Aristotelians and Stoics, whether plants had souls). Some things merely react; animals initiate action. But magnets too initiate motion; so they must be animals, and hence, definitionally, possess souls.

That argument is not of course compelling. It assumes without justification both that motion-causing is a basic, not a derived property, and that it is a causally unified concept (i.e. that there can only be one basic way in which motion is caused). But for all that, the argument is a singular and unprecedented piece of reasoning. Thales invites us to examine the bases of our metaphysical divisions in the world in order to uncover their causal and conceptual structure; and then to follow out their implications unshrinkingly even when they issue in striking and initially counterintuitive results. It is, in short, a paradigm of empiricism.

(d) Anaximander

Anaximander, like Thales, came from Miletus; and if Diogenes (*Lives* 2. 2, reporting Apollodorus) is correct, he lived from around 610 to 540 bc . He made the first map (Agathermus, *Geography* 1. 1 = 12 A 6 DK = 98 KRS), introduced the sundial into Greece (Diogenes, *Lives* 2. 1 = 94 KRS), and wrote a book, later called by the generic title *On Nature*, of which one fragment survives (**8** below). He speculated on the origins of the human species: they must have originated from other animals—fish, actually—since human infants require an extensive period of nurturing (ps.-Plutarch, *Miscellanies* 2. 908d = 12 A 10 DK = 134 KRS; cf. 135–7 KRS). The rudiments of scientific methodology are again apparent: first identify a problem (how can the human species have originated?); then suggest a hypothesis to account for it (they grew out of more self-sufficient creatures).

Anaximander's range is evident in the following doxographical report:

6(i) He said that some certain unlimited nature is the origin of things, from which are generated the heavens and the world in them. It is eternal and ageless, and encloses all the worlds. He speaks of time, generation and existence and destruction being determinate. (ii) He said that the unlimited (*apeiron*) is the origin (*archē*) and element (*stoicheion*) of things, and was the first to call it ‘*archē*’. Furthermore, there is an eternal motion, as a result of which the heavens are

end p.13

generated. (iii) The earth remains aloft, unsupported by anything, because of its equidistance from everything. It is rounded and circular in shape, like a stone pillar. We walk upon one of its surfaces, the other being opposite. (iv) The stars are generated as a circle of fire, separated off from the fire in the world and surrounded by air. There are certain flute-like passages, or blow-holes, through which the stars appear; hence eclipses occur when the blow-holes are blocked. (v) The moon appears sometimes to wax, sometimes to wane, according to whether the passages are blocked or open. The circle of the sun is twenty-seven times larger [13] of the moon , and the sun is highest, the circles of the fixed stars lowest. (vi) Animals are generated evaporated by the sun. Humans originally were like another type of animal, namely fish. (vii) Winds occur when the lightest vapours are separated off from the air, moving whenever they congregate; rain from vapour sent up from the earth by the sun; lightning when wind falls upon the clouds and tears them apart. (Hippolytus, *Refutation of All Heresies* 1. 6. 1–7 = 12 A 11 DK = 101b , 122b , 124–5, 129, 136 KRS)

Anaximander's conception of the structure of the heavens is striking and original: what appear as heavenly bodies are in fact the fires visible through the holes in hollow wheel-rims, and the fixed stars are the nearest of all (**6**(iv–v)). This ingenious theory further seeks to account for two of the more puzzling features of celestial geography: the phases of the moon and eclipses;[14] and **6**(vii) is the earliest attempt at serious, non-mythological meteorological explanation. Even so, Anaximander's cosmology is frequently arbitrarily dogmatic: how can he know that the circle of the sun is twenty-seven times greater than that of the earth, or that the earth is a cylinder three times as wide as it is high (ps.-Plutarch, *Miscellanies* 2 = 12 A 10 DK = 122 a KRS)?[15]

(e) The Principle of Sufficient Reason

The suggestion of **6**(iii) is amplified by Aristotle:

7some say that it [i.e. the earth] stays put because of similarity (thus, among the early philosophers, Anaximander). For there is no reason why what is situated in the middle and is similarly placed with regard to the extremities should move upwards rather than downwards or sideways. But it cannot move in opposite directions at the same time. So of necessity it stays where it is. (Aristotle, *On the Heavens* 2. 13. 195^{b}11–16 = 12 A 26 DK = 123 KRS)

end p.14

The world is a centrally positioned squat cylinder. Assume the innermost surface of the surrounding cosmos to be spherical; then for every line from the earth's centre to any point on that surface there will be an equal and opposite radius to be drawn in the other direction. Any motion must follow some radius; but there is no compelling reason for it to follow any one any more than its opposite: hence the earth will remain stationary at the centre.[16]

Anaximander's genius is to see how what appears to be a genuine physical problem (why does the earth not fall?) may, on a sufficiently acute and incisive examination of its presuppositions, cease to be puzzling at all. The earth does not fall (or indeed move in any way) because, given the structure of the heavens, there is no reason why it should: objects will only move in a suitably differentiated space, and observation shows us that our universe is not (in this sense at least) so differentiated. It seems to be a brute fact that objects fall downwards, but the idea that there is a single privileged direction for their travel is mistaken: we are misled by the phenomena as they appear to us, on the earth's surface. Anaximander anticipates Leibniz's 'Great Principle' of Sufficient Reason (PSR) (*Monadology* §32): if something is to occur, there must be a reason why it does so which is not outweighed by any countervailing force (although Anaximander presumably did not formulate it with any such generality). PSR, in effect, demands that events in the world have causes, and that, as such, they must be in principle explicable. Anaximander presumably relied upon everyday empirical evidence to come to that general conclusion; but it allows him to infer that, far from being remarkable and puzzling, the earth's motionlessness is only to be expected.

Furthermore, Anaximander's disposal of the problem implicitly accepts that explanation must be general. Its not moving is not due to any peculiar, intrinsic feature of the earth as such. Rather anything so placed in regard to the extremities of a uniform universe will behave the same way: if *a*'s being *F* explains its being *G*, then anything *F* must be *G*. Equally, explanations should at least aspire to completeness; if there is a genuine class of *G*-type things, then their individual *G*-ness must be susceptible of a unified explanation. Thus explanations should, ideally, be both necessary and sufficient for what they explain. That view is powerful in its complete generality, and seductive in its promise of a fully intelligible world. But ultimately its truth must be an empirical matter: there is nothing logically impossible about causal indeterminacy, and in an indeterministic world PSR fails to hold, at least in its full generality. But even so, to the extent to which they are explicable, the workings of the universe must conform to

PSR; PSR becomes, then, a regulative principle on the pursuit of scientific explanation (see VI.1*e*).

(f) The Unlimited

Anaximander talked of the *apeiron*, or 'unlimited', perhaps also describing it as an *archē*, or origin (**6**(i–ii)). *Archē*, which also means 'principle' or 'axiom', and derives from a word meaning both to govern and to begin, is to become a crucial component of the Greek vocabulary of explanation.[17] *Apeiron* is a negative noun, formed by prefixing the alpha of negation to *peras*, or limit.[18] Thus the *apeiron* is what is without limit, or limiting factors. In later Greek, *apeiron* means, literally, infinite in extent, and Aristotle interpreted Anaximander's *apeiron* in this way (*Physics* 3. 4. 203^{b}3–16) perhaps rightly: the notion of being without spatial limit involves no refined mathematical understanding of the infinite. On the other hand Anaximander's *apeiron* may rather have signified stuff without boundaries in the sense of being undifferentiated, formless.[19] Qualitative indifference well suits the role Anaximander assigns to his *apeiron*:

8(i) he says that it is neither water nor any of the other so-called elements but some different infinite nature, from which all the heavens and all the worlds are generated. (ii) And the things from which existing things are generated are also those into which they are destroyed, *in accordance with necessity, for they give justice (*dikē*) and restitution to each other in accordance with the arrangement of time* [12 B 1 DK] (he talks of them like this in poetical language). (iii) It is clear that he observed the change of the four elements into one another, and was unwilling to make anyone of them the underlying stuff (*hupokeimenon*) but rather chose something else apart from them. (iv) He accounts for their generation not by the alteration of the element but by the separating off (*apokrinesthai*) of the opposites by the eternal motion. (Simplicius, *On the 'Physics'* 1. 2. 24. 17–25 = 12 A 9 DK = 101 a KRS)

Anaximander might have felt that his basic material had to be actually infinite, or else the mechanisms of the universe will eventually run down (cf. Anaximenes, § *g*, **12** below). Yet if Simplicius' summary is accurate, Anaximander espouses a Principle of Conservation, of the sort which does away with the need for an unlimited mass of background material to replenish the processes of generation (cf. Aristotle, *Physics* 3. 8. 208^{a}8–11). Things change into one another endlessly, but without affecting the Universe's
end p.16

total mass. An eternal but finite world need never run out of causal steam; the 'eternal motion' of **6**(ii) and **8**(iv) will see to that. Such conservation principles, in prescribing the limits within which change may take place (and thus ruling out arbitrary and unruly change), are basic to the scientific picture of the universe as an intelligible whole.

On the other hand, Anaximander may have had another, more fundamental, reason for his *apeiron* to be infinite: *why* should the universe be here rather than there? And why should it be arbitrarily bounded? Aristotle alludes to such an argument in *Physics* 3. 4. 203^{b}25–8, a passage which contains hints of Anaximandrean paternity; if this is right, Anaximander anticipates arguments later deployed by the Atomists, also inspired by PSR, for an infinite Universe (VI.1*e*).

Scholars disagree about the actual extent of the fragment preserved in **8**(ii); but the notion of cosmic justice is certainly Anaximandrean.[20] In describing the passage as poetic in language, Theophrastus (whom Simplicius follows here) presumably took it metaphorically: the Universe does not really follow the dictates of some quasi-divine justice. Anaximander does not tell us how the great cosmic cycle is powered (beyond adverting to the 'eternal motion'). But this is not necessarily an explanatory deficiency: admitting that some things are beyond explanation may rather be straightforward realism. Moreover, Aristotle was to emphasize (V.1*a*) that explanation has to stop somewhere if we are to avoid infinite regress or circularity. If we adopt this not unattractive view, then there will be some fundamental facts about the universe which cannot themselves be explained just because they *are* fundamental. If this is right, then the 'eternal motion' that powers the system, and its cyclical nature, are such basic facts.

Still, we may ask *why* Anaximander thought the world had such a structure; and **8**(iii) gives us, at second hand, Theophrastus' answer. How securely Theophrastus' account is based upon Anaximander's text is unclear; but it motivates Anaximander's choice of the *apeiron* as basic principle. For if all determinate stuffs may transmute into one another (as **6**(vii) and **8**(iii) suggest), it is arbitrary to pick on one of them (water, as it might be) to be basic. What underlies those changes must, then, be distinct from and more fundamental than any of the basic but determinate stuffs of which the ordinary objects of the world are made; and even if Theophrastus is guilty of anachronistically reading back the structure of a later element theory into Anaximander's account, there is no reason for thinking this picture to be fundamentally misleading.

Precisely what it might involve depends upon the interpretation of **8**(iv). Anaximander's *apeiron* might be either (*a*) a smooth, undifferentiated,

uniform stuff (cf. §4*b* below), or (*b*) a *mélange*, a soup in which a variety of distinct ingredients jostle together in close contiguity; that is, it might resemble (in contemporary terms) either a chemical compound or a mixture. Aristotle favours (*b*):

9 some say the opposites are separated out (*ekkrinesthai*) from the One, being present in it, as Anaximander says . . . for the other things separate out from the mixture. (Aristotle, *Physics* 1. 4. 187^{a}20–3 = 12 A 9 DK = 118 KRS; cf. *Metaphysics* 12. 1. 1069^{b}20 ff. = 120 KRS)

The 'opposites' are suspended in the *apeiron*, and precipitated out of it. Aristotle's verb describing the process, *ekkrinesthai*, is subtly different from the *apokrinesthai* of **8**(iv), which suggests rather a process of distillation (and hence favours (*a*)). Anaximander may himself have been unclear about the relation between the *apeiron* and the elemental opposites; indeed he may not have seen that there was an issue here, since he may have lacked the conceptual equipment necessary to differentiate (*a*) from (*b*). But the evidence strongly suggests that Anaximander's *apeiron* was indeed smooth, formless, and undifferentiated, in line with interpretation (*a*) (cf. **10** below).

6(i) suggests a plurality of worlds: scholars disagree as to whether they exist concurrently or successively.[21] I remark only that PSR might support the contention that there are indefinitely many worlds at any given time. The worlds are generated from the *apeiron*; and the *apeiron* underlies the cyclical changes within the worlds. Anaximander may have been influenced by the evident cyclicality of the seasons (thus KRS 119), and if so, his view is a weak analogical induction. The fact that some processes are cyclical does not imply, or even suggest, that all of them must be. But the hypothesis of cyclical eternal motion has at least the advantage of explanatory economy. Anaximander apparently saw that physical explanation demanded more than merely a reductive account of the composition of the world; it needs to show too where the system gets its inexhaustible dynamism from. Linear processes tend, of their very natures, to some finite end; and hence it makes sense to assume that the fundamental forces that power an eternal universe are not linear in form (for this supposition, see further §3*c* below; III.2*a*; V.1*a*).

Anaximander was clearly, on any interpretation, a thinker of power and originality. He was commendably unafraid of bold intellectual leaps and striking hypotheses, and prepared to stand conventional wisdom on its head in the pursuit of a new understanding of the universe. He was perhaps the first to see (at least if interpretation (*a*) is correct) that a fully satisfactory ontology and metaphysics of change requires entities at more than one level (particularly if his ‘opposites’ are properties, such as hot, cold, wet, and dry, rather than substances).[22] Explanation will now take the form of showing how objects at the higher levels depend for their structure and properties upon their constitution from more fundamental entities. Moreover, Anaximander does not baulk at making the most basic stuff of all a purely inferred, theoretical entity, against which Aristotle's objection, that such a body should be perceptible to the senses (*Physics* 3. 5. 204^{b}32–6), seems peculiarly misplaced. Above all, his explanations aim at comprehensiveness and all-inclusiveness. They are, at least in this sense, paradigms of scientific explanation.

(g) Anaximenes: The Priority of Air

Anaximenes (fl. 550[23]) is usually[24] stigmatized as an unworthy successor to Anaximander, abandoning the latter's bold theoretical innovations in favour of more naïve modes of physical explanation:

10(i) Anaximenes . . . an associate of Anaximander, also says . . . that the underlying nature is one and unlimited, but not undefined (*aoriston*) as Anaximander said, but definite, saying that it is air. And it differs in respect of its substance (*ousia*) in rarefaction and density. (ii) Being rarefied fire is generated, being thickened wind, then cloud, then water, then earth, then stone, and the other things from these. He too makes motion eternal, through which change comes to be. (Theophrastus, in Simplicius, *On the ‘Physics’* 1. 2. 24. 26–25. 1 = 13 A 5 DK = 140 KRS; cf. **13**)

Some have taken **10**(ii) to betoken a complex physics, where the basic substances are generated as modifications of air, and everything else as compounds out of them; but Hippolytus' report (*Refutation* 1. 7. 1–3) contains no such suggestion, and there is no other evidence that Anaximenes saw the need for a physics of composition as well as one of density and rarefaction. Air is materially basic; and it is also internally dynamic, which accounts (somehow) for the condensations and rarefactions.

The decision to make air basic seems at first arbitrary and unmotivated. Since everything can be changed into everything else, it is hard to see why air should be somehow prior to water or stone. Yet Anaximenes clearly thought it was:

11Anaximenes held that air is the origin (*archē*) of existing things; for from it everything comes to be, and into it everything is resolved; *just as our souls*, he says, *being air, hold us together, so breath (*pneuma*) and air contain the whole world*[25] [13 B 2 DK]. (Aëtius 1. 3. 4 = 160 KRS)

Some have taken Anaximenes here to anticipate Plato's doctrine of the World-Soul (III.2*a* below). But it seems preferable to suppose him to be arguing that if air (being the substance of our souls) is responsible for our movements, then economy in explanation prompts us to consider the hypothesis that it is responsible for cosmic motions as well (although if construed as an argument for the thesis, **11** is a spectacularly feeble induction).

But he must still establish that our souls really are air. Another text, besides reaffirming the unity of the motive cause, hints how he did so:

12 Anaximenes believes that there is one unlimited motive principle (*archē*) of all things, namely air. For he says as follows: *air is close to the incorporeal; and because we come into being by an outpouring of air, it is necessary that it be both unlimited and rich on account of its never giving out*[26] [13 B 3 DK]. (ps.-Olympiodorus, *On the Sacred Art of the Philosopher's Stone* 25)

The basic stuff from which everything else is generated must be unlimited, otherwise it will eventually give out; but, if the universe has had no beginning and the basic stuff were finite, it would already have done so (cf. III.2*a*; V.1*a*). Anaximenes steps back from Anaximander's conservation principle.

Anaximenes wishes to assimilate the air we breathe to the basic force of life itself; and of course a steady supply of air is necessary for the preservation of life. But it is not sufficient; and it is not yet clear why air should be exalted over any one of a number of other equally essential prerequisites. However, **12** suggests another reason for doing so: air 'is close to the incorporeal', i.e. to the insubstantial. Anaximenes is talking not of the vaporous *aēr*, but of the invisible atmospheric stuff which, while having perceptual effects, still seems remote from human contact by comparison with any other candidate element. Air will then approximate to an Anaximandrean *apeiron*, ready to take on qualities as a result of condensation and rarefaction, but in its airy state actually possessing none of them:

end p.20

13 the form of air is thus: whenever it is most even it is invisible to sight, but is revealed by the cold and the hot and the moist and by movement. It is always in motion; for the things that change do not change unless they are moved. Through becoming denser or finer it has different appearances . . . [Here Hippolytus describes the same process as **10**.] Consequently the principal components of generation are opposites, hot and cold. (Hippolytus, *Refutation* 1. 7. 1–3 = 13 A 7 DK = 141 KRS)

(h) Analogy and Explanation

Anaximenes was fond of analogies (as **11** shows). His view that the earth rides on air seems at first sight fairly hopeless: earth clearly falls through air, and air is far too insubstantial to support gross earth. But:

14 Anaximenes and Anaxagoras and Democritus say that its [i.e. earth's] flatness is the cause of its staying still, since it does not cut the air beneath, but covers it like a lid, as flat bodies are seen to do; for they are difficult to move even for the winds because of their resistance. They say that because of its flatness the earth does the same in relation to the air under it (which not having sufficient room to move away, stays motionless in a mass below) like the water in a pipette. (Aristotle, *On the Heavens* 2. 13. 294^b13–21 = 13 A 20 DK = 150 KRS)

The illustration of the saucepanlid supported by boiling vapour does not sit particularly well with Anaximenes' conception of the earth completely surrounded by air.[27] Yet the fact that heavy lids can be supported by insubstantial steam does perhaps show that there is nothing a priori absurd in the notion of the earth's resting on air.

Equally, the pipette analogy is not exact: water does not stay in the pipette because it has nowhere to go. Equally, the earth has nowhere to go in displacing the air, not because the air is contained within something, but because air exhausts all the infinite space there is. If the earth were to fall, it would push air before it, which in turn would displace more air: and so on. But the universe being a plenum, there is no empty space for the displaced air to move into; hence the earth cannot move. That argument fails, since it is false (at least on non-corpuscularian assumptions) that the air would have nowhere to go in a plenum, since it may move reciprocally, filling the vacant space left behind the moving body, thus requiring no actual void.[28]

Anaximenes has no apparent need of physical heavenly spheres in his infinite airy universe: the earth rides on air because (partly) of its shape,

while it is because the sun is flat 'like a leaf' (Aëtius 2. 22. 1 = 155 KRS) that it stays suspended aloft (Hippolytus, *Refutation* 1. 7. 4 = 151 KRS). But Aëtius also says (2. 13. 10 = 152 KRS) that among Anaximenes' visible, fiery heavenly bodies are mixed invisible earthy ones, which might suggest that the visible ones are embedded in earth (these invisible earthy bodies may have been designed to account for eclipses, like Anaximander's 'wheel-rims': above, § *d*). Moreover,

15 Anaximenes said that the stars are fixed into the crystalline like nails. (Aëtius 2. 14. 3 = 13 A 14 DK = 154 KRS)

This need not commit Anaximenes to a physical system of celestial crystal spheres.[29] Analogies may illuminate by suggesting a physical mechanism for some puzzling fact by extension from some other, less intrinsically baffling phenomenon. But more weakly, they may merely suggest that the surprising fact is sufficiently similar to something familiar for our puzzlement at it to be reduced; and **15** should probably be interpreted thus. The stars are not nailed to the heavens; but if we wonder how they can possibly stay up, we need only consider the nails in a ceiling to see that such apparent suspensions are possible.[30]

In regard to thunder and lightning,

16 Anaximenes says the same as him [i.e. Anaximander], adding the phenomenon we observe on the sea, which gleams when cut by oars. (id. 3. 3. 2 = 13 A 17 DK = 158 KRS; cf. Hippolytus, *Refutation* 1. 7. 8)

This too may be no more than an appeal to one phenomenon to explain something superficially similar, albeit on a much larger scale, and thereby to reduce our tendency to think of lightning as supernatural, and hence fearful. But Anaximenes may also mean to suggest that the tearing or parting of substances tends to produce luminescence, and thus to imply the possibility of an explanatory generalization covering two apparently quite diverse cases.

This certainly appears to be the strategy in the case of another terrifying phenomenon:

17 earthquakes are caused when the earth is greatly altered by heating and cooling. (Hippolytus, *Refutation* 1. 7. 8 = 13 A 7 DK)

18 Anaximenes says that the earth, as a result of being soaked and then dried out, fractures, and is shaken by those masses which are torn off and fall in. Consequently earthquakes occur in periods both of drought and of heavy rain; for in drought, as was said, it dries up and fractures, and then falls apart when saturated with water. (Aristotle, *Meteorology* 2. 7. 365^{b}6–12 = 13 A 21 DK = 159 KRS)

19 Just as in old buildings certain parts give way although not struck, when they have more weight than strength; so in the earth as a whole it occurs that certain parts are loosened by age, and being loosened fall and cause the parts above them to tremble. (Seneca, *Natural Questions* 6. 10)

17 and **18** are almost certainly Anaximenean; and **19** has been plausibly ascribed to him. Earthquakes are massive physical events which occur apparently without reason or warning, and thus invite supernatural explanation. Anaximenes' attempt to strip them of their terrifying numinous aura is paradigmatic of the new Milesian rationalism.

First **17** proposes a normal, physical cause for earthquakes: heating and cooling. **18** builds upon that by suggesting complementary mechanisms with which excessive dryness and moisture may be responsible for the structural damage that precedes a quake. Finally, **19** shows how that damage may be cumulative, and hence why there may be no particularly catastrophic meteorological precedent for an earthquake (and why they appear to strike at random). Further, **18** and **19** develop a picture of the earth's internal structure in virtue of which it can behave in this way: it is hollow, with internal peaks and crags which are subject to weathering just like their superterranean counterparts. Thus earthquakes are not (as they seem) without physical cause. Their occasions may appear trivial, perhaps even non-existent (as is sometimes the case in structural collapse): but for all that they are susceptible of a complete physical explanation.

What emerges is a clear hypothesis of great explanatory power. It is suggested by analogy, but the analogy does not simply take the form of a feeble induction. Rather it shows how normal (if imperfectly understood) physical processes may be invoked as causal hypotheses to explain events with which they have no obvious, superficial similarity. And that is one of the hallmarks of bold, conjectural science.

(i) The Reduction of Properties

For Anaximenes, different types of stuff are reduced to functions of the denseness and rarefaction of the basic air. But he goes a stage further:

20 (i) should we, as old Anaximenes supposed, treat the hot and the cold not as substances but rather as common properties of matter which supervene on changes? (ii) For he says that matter which is concentrated and condensed is cold, while that which is rare and *slack* (that is the word he uses) is hot. (iii) Hence not improbably are men said to exhale both hot and cold from their mouths; for the breath is cooled when it is compressed and condensed by the lips, but when the mouth is relaxed and it is exhaled it becomes hot by reason of its rarefaction. (Plutarch, *On the Primary Cold* 7. 947f–8a = 13 B 1 DK = 143 KRS)

In addition to prefiguring the later debate on whether hot and cold are properly substances or properties, **20**(i) poses a sophisticated question about their metaphysical status. Many later Greek physical systems (including Aristotle's) made them basic; Anaximenes, according to **20**(ii), believed hot and cold to be derived properties, supervenient upon the density and rarefaction of the stuff embodying them (and, first appearances notwithstanding, Hippolytus in **13** may simply mean that heat and cold are the principal phenomenal *signs* of rarity and density for Anaximenes).

This view is not without its difficulties (why don't substances always exhibit the temperature appropriate to their specific densities, making hot stones and cold winds impossible?). But the distinction between basic and derivative properties is original, profound, and still fundamental to physics: heat just *is* mean molecular kinetic energy, which is itself a function of density and rarefaction. But if heat and rarefaction are identical, why should one of them be prior to the other? In cases where a variety of apparently distinct *x*s are all found at bottom to be *y*, then reductive simplicity justifies our making *y* basic; but here there seems no such warrant. The obvious thing to say is that the priority in question is causal: if *x* really is *y* (for coextensive *x* and *y*), then *y* is causally responsible for *x*. Density causes coldness, rather than vice versa. This is supported by experiment: it is by making the air dense (by pursing my lips) that I make it cold—I do not first chill the air, causing my lips to purse. The direction of the intervention indicates the direction of causal dependence.[31]

Thales, Anaximander, and Anaximenes emerge from our examination with their revolutionary reputations intact. Their observations were patchy and inadequate; but at least they saw the relevance of observation to explanation. They did not, moreover, fight shy of postulating, on theoretical grounds, the existence of objects which were not directly observable. And, although prone to the intoxication of the audacious, sweeping conjecture that so impressed Popper (1958), they saw that genuine explanation of the world and its processes can only be won by way of large and

simple general theories that seek to comprehend the bewildering variety of physical phenomena within their grasp.

(j) Xenophanes: Elemental Explanation

Xenophanes of Colophon in Ionia probably lived from around 580 to 480 bc . He was a poet, who wrote incidentally on philosophical matters. He solved the problem of the earth's rest by supposing us to inhabit the upper surface of an infinitely extensive mass of earth (21 B 28 DK = 180 KRS). Both earth and water are basic:

21 earth and water are all the things that come to be and grow (Sextus Empiricus, *Against the Professors* 10. 34 = 21 B 29 DK = 181 KRS; cf. 21 B 33 DK = 182 KRS)

21 earth and water are all the things that come to be and grow (Sextus Empiricus, *Against the Professors* 10. 34 = 21 B 29 DK = 181 KRS; cf. 21 B 33 DK = 182 KRS)

22 sea is the source of water, and the source of wind; for neither [32] from inside without the great ocean, nor river-streams, nor showery water from the *aithēr*; but the great ocean is the begetter of clouds and winds and rivers. (Geneva scholiast on the *Iliad* 21. 96 = 21 B 30 DK = 183 KRS)

The heavenly bodies were either particles of fire coalesced out of clouds, or ignited clouds (175–7 KRS), and rainbows were forms of cloud as well (21 B 32 DK = 178 KRS). The details of much of this are obscure, and the astronomy, in particular, is bizarre. There are, apparently, innumerably many suns, since the sun is every day a new object kindled from sparks, or clouds, in the east, and moving in a straight line to the west, and consequently the apparent curvature of its path is, according to Xenophanes, an optical illusion (Aëtius 1. 14. 3 = 175 KRS; Aëtius even suggests that there are different suns for different regions of the earth, although this is probably a mistake: 2. 24. 9 => 179 KRS). But it emerges moderately clearly from the prehistoric murk that Xenophanes adopted a two-stage cosmology, of the type some have ascribed to Anaximenes (§ *g*, **10**(ii) above). At the bottom are the elements; but elaborated from them are further compound substances, one at least of which (cloud) is further susceptible of modification into different forms. Xenophanes may have thought that rainbows were *really* cloud (and not water), although cloud itself is elaborated water; or he may have held more simply that becoming cloud is an intermediate stage in water's transition to rainbow. But however that may be, Xenophanes' cosmology and meteorology are new developments in the notion of reductive explanation.[33]

end p.25

3. Causal Origination

(a) *Xenophanes: Reason and Religion*

The Milesians were optimistic, reckless even, in their assessment of the possibility of human knowledge. Xenophanes, by contrast, openly admitted that knowledge of nature was difficult, perhaps impossible, to come by:[34]

23 The clear truth (*to saphes*) no-one has ever known, nor will know, Concerning the gods and all the things of which I speak; For if he should happen in fact to utter the whole truth, He himself will yet not know: belief reigns over all. (Sextus Empiricus, *Against the Professors* 7. 49, 110 = 21 B 34 DK = 186 KRS)

However, he allows elsewhere (21 B 18 DK) that, with practice and diligence, human beings may make modest progress in discovery.

But it is Xenophanes' theology which is of particular concern to us. Thales thought the world full of gods (§2. *b*, **4** above); while in more abstract vein Anaximenes considered his all-encompassing air to be divine (141, 144–6 KRS). Xenophanes carries the theological revolution further with a powerful assault, founded upon comparative anthropology, on the traditional, anthropomorphic conception of the gods:

24 the Ethiopians say that their gods are black and snub-nosed, while the Thracians say that theirs are blue-eyed and red-haired; (Clement, *Miscellanies* 7. 22. 1 = 21 B 16 DK = 168 KRS)

consequently,

25 if cows, horses, and lions had hands, and were able to draw with their hands and do the work men do, horses would draw images of gods like horses and cattle like cattle. (ibid. 5. 109. 3 = 21 B 15 DK = 169 KRS)

The gods of myth are simply humans writ large and unpleasant:

26 Homer and Hesiod attributed to the gods all the things which among men are shameful and blameworthy—theft, adultery, and deception of one another. (Sextus Empiricus, *Against the Professors* 9. 193 = 21 B 11 DK = 166 KRS)

Xenophanes is a debunker of popular religion, offering a 'natural history' for ordinary beliefs which shows how men could have come to hold them in spite of their falsity. The epic poets of **26** are mistaken: gods simply could not be like that. But Xenophanes is no atheist. His positive theology survives largely in a Theophrastean passage of Simplicius (*On the*

'Physics' 1. 2. 22. 26–23. 20). His argument starts with a truth of conceptual analysis:

(1) God is the most powerful of all things,

and so

(2) God is unique:

27 if there were more than one, he says, they would all have to have equal power, but what is most powerful and best of all things is God. (ibid. 1. 2. 22. 31–3 = 21 A 31 (3) DK)

Next Xenophanes

28 shows that (3) God is ungenerated, from the fact that (4) what comes to be must do so either (*a*) from what is similar or (*b*) from what is dissimilar, but (5) similar things cannot be affected by one another (for it is no more appropriate that what is similar should generate than that it should be generated by what is similar to it); and if it comes from what is dissimilar, then (6) what is will be from what is not. In this way he showed God to be ungenerated and eternal. (1. 2. 22. 33–23. 4 = 21 A 31 (4) DK)

Assume the negation of (3); then either (4*a*) or (4*b*) must be true; but neither are; so, by *reductio*, (3). (6) is taken to be conceptually impossible: nothing can come to be from nothing (see § *d*, CP1 below). But why should (6) tell against (4*b*)? (4*b*) may yield

(4*b**) *x* comes to be *F* from not being *F*;

and if we allow existence to be a predicate, then x comes to exist from not having existed, which is supposedly impossible. It is not at first sight clear why: but Xenophanes may have supposed that (4*b**) must mean that the non-existence could in some strong sense *produce* something, which is impossible, since what does not exist cannot bring about something that does. The non-existent can have no causal powers at all. Versions of that argument, championed by Parmenides (§ *d* below), reverberate down the ancient world.

Equally Parmenidean is the dilemma's other horn. Similar things cannot bring about similar things; *x*'s *F*-ness cannot be the cause of *x*'s coming to be *F*, for if it were *x* would have been *F* all along, and hence there would be no coming to be. The argument as Simplicius presents it employs language that cannot be Xenophanean, and some have thought it represents a later, mistaken reconstruction of Xenophanes' thought in the light of Parmenides.[35] Xenophanes may rather have argued that if God were
end p.27

created, it must have been by some power; but not by any power stronger than himself, since by (1) there are none; but nor yet by any weaker power, since the weaker cannot create the stronger; finally no equal power will do, since, on the strong construal of (1) represented by (1*a*), there are no equal powers either. Equally, nothing could destroy God unless it were, *per impossibile*, more powerful.

But that Cartesian argument, if indeed it is to be attributed to Xenophanes,[36] is hardly compelling, unless existence can be squeezed out of conceptual analysis, since all it shows is that *if* something satisfies the concept of the divine, *then* it must be immortal and ungenerated; and perhaps Xenophanes intended it to do no more. But theologians since Anselm have often sought to conjure existential rabbits from essential hats, and Xenophanes may have attempted a similar feat of metaphysical prestidigitation. But whatever its faults, Xenophanes' argument represents the first attempt at rational theology, and it invokes abstract considerations concerning the nature of causing and generation that are to be of fundamental importance to subsequent Greek thought. We are already far from the Rabelaisian world of Homer's Olympians.

(b) Heraclitus

Heraclitus of Ephesus, who flourished around 500 bc , and whose writings became a byword even in antiquity for their obscurity, inaugurated a sharp new metaphysical turn in Presocratic philosophy; but he was in many respects a typical early Presocratic. The heavenly bodies are for him hollow bowls filled with fire (Diogenes, *Lives* 9. 9–10 = 224 KRS), eclipses and phases of the moon being caused when the bowl turns its back towards us. Things are cyclical in nature (‘beginning and end are common’: 22 B 103 DK), and there is a general interchange of elemental opposites, in the manner sketched by Anaximander (‘cold things grow hot, the hot cools, the wet dries, the parched moistens’: 22 B 126 DK; cf. B 31, 91 DK), and under a similar general divine ordinance of conservation. The soul too is subject to transmutation:

29 it is death for souls to become water, death for water to become earth; but from earth water is generated, and from water soul. (Clement, *Miscellanies* 6. 17. 2 = 22 B 36 DK = 229 KRS)

Soul is fiery (22 B 36, 45, 117–18 DK), and we become sentient by absorbing divine fire from the surrounding atmosphere (Sextus Empiricus, *Against the Professors* 7. 129), fire being the dominant element into which everything is eventually resolved (**31–2**).
end p.28

Heraclitus apparently subscribed to two general metaphysical hypotheses:
(H1) there is a constant dynamic motion in things (the 'Theory of Flux');[37]
and
(H2) all things are compositions out of opposites.
The flux of H1 indeed accounts for what permanence ordinary objects possess:

30 things which have a natural circular motion are preserved and stay together because of it; if indeed, as Heraclitus says, the barley-drink separates if it is not stirred. (Theophrastus, *On Vertigo* 9 = 22 B 125 DK)

This idea, that stability is in fact owed to a sort of dynamic equilibrium, was to be taken up by the Stoics (VII.1*b*) and Numenius (XII.1*b*). H2 was apparently to be interpreted in a very strong sense: opposites are constantly coexistent in everything; and each thing perhaps even is its opposite (22 B 10, 51, 60–1, 67, 88, 101, 111 DK), a position Aristotle was later to castigate as straightforwardly incoherent (*Metaphysics* 4. 3. 1005^{b}23–6). Aristotle for one held that the belief that things really are opposites (which he also attributes to Anaxagoras and Democritus) derives from perplexity arising from the notion of generation: people see contrary properties coming to be in the same thing, and assume therefrom that the contraries must in some way constantly be present in it (ibid. 4. 5. 1009^{a}23–7), otherwise something will have come to be, absurdly, from nothing (cf. § *d*, CP1 below).

Such concerns appear particularly sharply after Parmenides and the Eleatics explode their mine under the conceptual structure of Presocratic physics (§§ *d–e* below). But they may already have pushed Heraclitus to adopt a particularly radical version of H2. And, taken together, H1 and H2 do show how change, generation, and alteration are not merely possible, but continuous features of the world we live in, which was, consequently, never the product of a cosmogony:

31 neither did any god or man make this world, but it was always and is and shall be: an ever-living fire, catching alight in measures and extinguished in measures. (Clement, *Miscellanies* 5. 104. 1 = 22 B 30 DK = 217 KRS; cf. 22 B 31 DK)

end p.29

The notion of measure, as well as the continuous and eternal nature of the process, recalls Anaximander's cosmic justice (§2*f*, **8** above), as do other fragments:

32 everything is exchanged in return for fire and fire in return for everything, as goods are for gold and gold for goods. (Plutarch, *On the 'E' at Delphi* 8. 388d = 22 B 90 DK = 219 KRS)

Compare 22 B 94 DK: 'the sun will not overstep its boundaries, or else the ministers of justice will find it out'. The sun is 'new every day' (22 B6 DK), as it apparently was for Xenophanes (§2*j* above), created from exhalations of the sea as part of the eternal cycle of transmutations; and 'justice' sees to it that balance, order, and regularity are maintained (22 B 94 DK).

Finally, we may discern a distinctive new note of divine teleology: the 'thunderbolt' is said to organize, or steer everything (22 B 64 DK), while

33 to God all things are fair and just: but men thought some just, others unjust. (Porphyry, *On the 'Iliad'* 4. 4 = 22 B 102 DK = 206 KRS)

That latter remark prefigures, if obscurely, Stoic theodicy (VII.3*a*); and the Stoics are indebted to Heraclitus' physics and metaphysics in other ways as well. For all its legendary obscurity, Heraclitus' philosophy presents the world as one composed of intelligible processes which human beings can (in principle) come to understand. Moreover, his world is intelligible because, like Anaximander's, it is fundamentally governed by law and reason, or *logos*: an underlying structure which will account for and make comprehensible the world of ordinary experience, and which the diligent inquirer (22 B 101 DK) can ultimately make patent, even if his pronouncements will not be understood, let alone believed, by the vast, ignorant, gullible majority (22 B 1–2 DK; cf. B 17, 19, 28, 34, 56, 72, 78, 107 DK). Heraclitus is thus the first thinker in the Western tradition to adumbrate a physics in which the ultimate microstructure of the world is quite distinct from anything immediately apparent to the senses, and to which we must infer by reason, albeit reason guided by and responsive to the way things appear (compare the physics and epistemology of Anaxagoras, §4*b* above; and Democritus, VI.1).

(c) Alcmaeon and The Soul

Alcmaeon of Croton in Italy (fl. *c*.480 bc), a doctor who was connected with the Pythagoreans,[38] propounded an empiricist theory of perception

end p.30

and cognition,[39] as well as formulating a philosophical account of health in terms of the balance (or equity: *isonomia*) of opposites (Aëtius 5. 30. 1 = 24 B 4 DK). He was also concerned with natural philosophy, averring, obscurely, that 'most human affairs are dual' (Diogenes, *Lives* 8. 83 = 24 A 1 DK), and remarking that

34 human beings die because they cannot join the beginning to the end. (ps.- Aristotle, *Problems* 17. 3. 916^{a}33–4 = 24 B 2 DK; cf. Heraclitus 22 B 103 DK above).

'Joining beginning to end' suggests completing a circle; and **34** suggests that human life, being a linear process, must necessarily terminate. The same does not, however, go for the human soul:

35 he says that it is immortal on account of its similarity to the immortals; and this is the case for it because it is always in motion; for all the divine things move continuously, moon, sun, the stars and the whole heaven. (Aristotle, *On the Soul* 1. 2. 405^{a}30–^{b}1 = 453 KRS = 24 A 12 DK)

36 Alcmaeon takes the soul to have a self-moved nature in eternal motion, and for this reason it is immortal and similar to the gods. (Aëtius 4. 2. 2 = 454 KRS = 24 A 12 DK)

Aristotle makes Alcmaeon's argument an unimpressive analogy: the soul always moves, and the heavens always move, so the soul is like the heavens; but the heavens are immortal, so the soul is immortal. But Aëtius implies rather that Alcmaeon argued directly from the soul's eternal motion to its immortality:

(1) soul is the cause of movement;

and hence

(2) the soul is itself in motion.

But given that

(3) the soul cannot cease moving,

then
(4) the soul is immortal.
(1) entails (2) by way of what is sometimes called the Principle of Causal Synonymy (PCS).[40] PCS may be formulated in different ways: but the intuitive idea is that if an agent *a* is causally responsible for some property
end p.31

F holding of *b*, then *a* must itself possess *F*, and make *b F* in virtue of its *F*-ness; thus motion-causers must themselves move. The plausibility or otherwise of various forms of PCS will loom large throughout the rest of this study. But if PCS motivates Alcmaeon's move from (1) to (2), it is still unclear why he accepts (3). He may have argued that
(3*a*) (i) if it is the function of the soul to cause motion, then it is essential that it move;
(ii) if it is essential that it moves, then it moves of necessity;
(iii) if it moves of necessity, then it cannot cease moving;
or alternatively,
(3*b*) (i) psychic motion is causally autonomous (i.e. unaffected by external factors);
but
(ii) if (i), then nothing external can interfere with it such as to stop it;
hence
(iii) its motion is eternal.
(3*a*) involves a modal scope-fallacy, of the sort generally assumed to vitiate Parmenides' fundamental argument (§ *d* below). But (3*b*) is, I think, more philosophically interesting, and raises profound questions in the analysis of causal origination. Animals have the power of self-movement; and it is a Thalean conceptual truth that they do so in virtue of their souls. Their souls are thus responsible for their movement, and hence (by PCS) themselves in motion. But, given that (*a*) animal motion is self-caused, then (*b*) souls must *originate* movement; and if (*b*), then (*c*) they are causally autonomous with respect to motion, i.e. (*d*) nothing can affect their motion-causing abilities. But in that case nothing can destroy (or for that matter create) them; hence if they exist at all they do so for ever.
That argument is not compelling. First, it is not clear that (*c*), in any sense in which it follows from (*b*), ought to be glossed as (*d*). To be causally autonomous means to have the power to initiate sequences of events, i.e. there are some sequences in which a soul is the first cause, and there is nothing prior to the action of the soul *in the sequence*; but that does not show that there is nothing causally prior to the soul *tout court*. Secondly, it is not clear, ordinary intuitions notwithstanding, that animals are in any case genuine initiators of causal sequences. Perhaps all acts, including acts of the will, are themselves antecedently caused, and agency is not to be differentiated from other types of causal activity in that it involves fresh causal starts (see VII.3–4). But even if fallacious, Alcmaeon's reasoning is
end p.32

subtle and sophisticated; at all events, Plato was sufficiently impressed to adopt it for himself (*Phaedrus* 245c–246a: III.2*a*).[41]

(d) Parmenides

Parmenides was born in Elea, a southern Italian Greek city, around 515 bc. A substantial portion of his philosophical poem survives, and we can, for once, base our interpretations largely upon his actual words. The poem fell into two main sections, the so-called Way of Truth and the much longer (although much less completely preserved) Way of Opinion,[42] preceded by a prologue in which Parmenides tells how a goddess takes him on a spiritual journey, promising to show him the real nature of things as well as demonstrating how humankind could have come to get things so badly wrong (28 B 1 DK = 288 KRS). She then says:

37 Come, and I will tell you (and you having heard must preserve the story)
What are the only conceivable roads of inquiry:
One, that it is and cannot not be,
Is the path of persuasion (for it attends upon the truth)
Another that it is not and must not be—
That I say to you is a trail devoid of all understanding:
For you could neither recognize what is not
Nor speak of it: for the same thing is for thinking as being.
(Proclus, *On the 'Timaeus'* 1. 345. 18–25; Simplicius, *On the 'Physics'* 1. 3. 116. 28–117. 1; Plotinus 5. 1. 8 = 28 B 2–3 DK = 291–2 KRS)

Two alternatives are offered:
(1) it is (and cannot not be)
and
(2) it is not (and must not be).
The verbs in (1) and (2) lack subjects. Owen (1960) plausibly supplies 'whatever can be thought of' as implicit subject,[43] as suggested by the last line of **37**, and by another fragment:

38 (i) What is to be spoken and thought of must be; (ii) since it can be
While nothing cannot.
(Simplicius, *On the 'Physics'* 1. 3. 86. 27–8 = 28 B 6.1–2 DK = 293 KRS)

A second crux concerns the predicate 'is' of (1) and (2). It may be existential, and thus (1) glosses as
end p.33

(1*a*) it (i.e. what can be thought or spoken of) exists (and cannot fail to exist);
or it may, in a standard Greek usage, be predicative, standing proxy for any predicate you like: i.e. for some value of F,
(1*b*) it is F (and cannot fail to be F).

Scholarly controversy has again been fierce;[44] I favour (1*b*) (of which (1*a*) may be seen as a special case, if existence is a predicate), although I shall continue to speak in existential terms for style's sake.

The modal parentheses are equally problematic. Parmenides may simply have meant that you cannot think of non-existent objects: what you conceive must exist. That ordinary-language sentence is, however, ambiguous between

(1*c*) it is necessary that whatever can be thought of exists

and

(1*d*) whatever can be thought of is a necessary existent.

Parmenides is often thought to have confounded (1*c*) and (1*d*), relying on (1*c*)'s plausibility to conclude fallaciously that nothing could ever be created or destroyed since everything is a necessary existent after the manner of (1*d*) (cf. § *c*, (3*a*i–iii) above). Furthermore, it is generally assumed that Parmenides' argument relies primarily upon logical considerations concerning the status of non-being. I want to challenge these orthodoxies.

38(ii)is supposed to support **38**(i); but it is unclear how, in default of reliance upon a modal axiom of staggering power and implausibility,

(3) *x* can be

is supposed to entail (or even support)

(4) *x* must be.

(3) is supplemented by a corollary:

(5) nothing cannot be,

i.e. presumably

(5*a*) the non-existent cannot exist.

(5*a*) is generally taken to be the locus of the modal fallacy, since it might be glossed either as

end p.34

(5*a*) (i) it is not possible that the non-existent exists[45]

or

(5*a*) (ii) what does not exist is not a possible existent.[46]

(5*a*i) is a trivial truth; (5*a*ii) generates the argument's logical leverage. From the confusion thus created, Parmenides steals the condusion that the only things which exist do so necessarily, and hence (via a standard ancient inference) eternally.

Yet I am inclined to acquit Parmenides of that logical felony. Rather he *argued* for the strong modal conclusion that what exists must do so eternally. In the next fragment he declares:

39for never will this prevail, that things which are not are, (Plato, *Sophist* 242a = 28 B 7.1 DK = 294 KRS)

which is usually taken to show that Parmenides' fundamental concern is with the logic of non-being; any proposition involving negation is meaningless, since all talk of non-being is incoherent. But **39** may rather make the claim that it is *physically* impossible for what does not exist to be generated, an interpretation supported by the following lines from the long fragment 8:

40 It never was nor will be, since it is all now together
One, continuous. For what coming-to-be will you seek for it?
How and whence might it grow? From not being I will not allow
You to say or to think. For it is not to be said or thought
That it is not. And what need would have impelled it
Later or earlier to be, beginning from nothing?
Thus it must either be completely or not at all.
Nor ever will strong credibility say that from not being
Can come to be something apart from itself.
(Simplicius, *On the 'Physics'* 1. 3. 145. 5–13 = 28 B 8.5–13 DK = 296 KRS)

40.6–8 (supported by 12–13) provides the fundamental reason why what is must be. If *x* comes to be *F*, then it must come to be from not being *F* (compare **28** above), otherwise trivially it will not have *come to be F* at all. But it cannot come to be *F* from not being *F*, since non-*F*-ness cannot causally explain its being *F*; the non-existent, being nothing at all, can have no properties, and hence *a fortiori* no causal properties. Only if the nonexistent could exist in some way (i.e. only if (5) were false) would the story begin to make sense. But if (5) is true, then only the existent can exist (i.e. (4) is true); hence what can exist must do so.

end p.35

Thus (4) is not meant, bizarrely, to entail (3). Rather the argument from generation will show that any time (4) is satisfied, (3) will be satisfied too. Thus the would-be generator is impaled on the horns of a dilemma deriving from causal considerations about what is rationally required for something to come to be, and which entails that generation is impossible. As a corollary, time too is shown to be non-existent, since time presupposes change.

That argument disposes of there being any causal explanation of *x*'s coming to be *F* (for any *F*—and hence of *x*'s coming to be at all); but might not *x* come to be *F* spontaneously? **40**.9–10 seeks to eliminate that possibility. Suppose (*per impossibile*) that *x* comes to be *F*, and hence comes to be *F* from not being *F*. Since not being *F* cannot explain *x*'s subsequent *F*-ness, there can be no *reason* why *x* comes to be *F*; but if there is no reason why *x* comes to be *F*, then there is no reason why it should do so at *t* rather than at *t*′; hence, by PSR, if it comes to be *F* at *t*, it should do so at *t*′ as well; but it is impossible for it to become *F* at both *t* and *t*′;[47] hence it cannot come to be *F*.

On this (admittedly unorthodox) account, Parmenides seeks to demonstrate the incoherence of change not on the basis of logic and semantics, but because of what he takes to be an incoherence in the notion of causality on the one hand, and the inexplicability of uncaused change on the other. Neither half of the argument is successful. It is question-begging to invoke PSR in order to claim that there can be no uncaused changes, since PSR is at best an empirical hypothesis. On the other hand, a suitably sophisticated analysis of the logical form of change, one which pays particular attention to the logically protean preposition 'from' in propositions such as '*x* comes to be from *y*', will dispose of the Parmenidean bomb; but it needed an Aristotle to defuse it (see IV.1*b*).

Parmenides' basic principle,
(CP1) nothing comes to be from nothing,
was to achieve canonical status. CP1, Hume's 'impious maxim', is ambiguous between
(CP1*a*) nothing comes to be causelessly,
and
(CP1*b*) nothing comes to be except from pre-existing matter;
and that ambiguity is not always patent. Indeed, distinguishing CP1*a* from CP1*b* is the first step towards solving the Eleatic puzzle, as Aristotle was to
end p.36

realize (IV.1*b*, **155**). That principle is not, however, original to Parmenides. It was ascribed to Xenophanes (**28** above), and occurs in a fragment of the sixth-century lyric poet Alcaeus, although we do not know in what context.[48]

Neither CP1*a* nor CP1*b* can be accepted a priori (except by stipulating that generation be defined as caused or materially grounded generation). But even so, there is a powerful reason for hoping that they are at least broadly true, or true for well-defined domains: for if they are not, the grand project of rendering the world intelligible by uncovering its mechanisms and regularity is condemned at least to partial failure.[49]

Only scattered fragments survive of the Way of Opinion. It purports to offer a typical Presocratic cosmology (the favoured elements being light and darkness: 28 B 9 DK = 303 KRS). Although Parmenides probably produces it solely exempli gratia as typical of fallacious mortal opinion (see n. 42), its astronomy is of some significance. The heavenly bodies consist of 'garlands', or rings, of fire (28 B 12 DK = 306 KRS); and he promises that

41 You will know both ether's nature, and all the signs
In the ether, and the destructive works of the pure
Torch of the resplendent sun, and whence they came to be,
And you will learn the wandering works of the round-eyed moon
And her nature, and you will know the encircling heaven,
Whence it grew, and how necessity led and shackled it
To hold the limits of the stars.
(Clement, *Miscellanies* 5. 138. 1 = 28 B 10 DK = 305 KRS)

The details of his account are lost; but two resonant lines survive concerning the moon:

42 A night-shining alien light, wandering around the earth, Always gazing towards the rays of the sun. (Plutarch, *Against Colotes* 1116a; *On the Face in the Moon* 929a = 28 B 14, 15 DK = 308 KRS)

Thus Parmenides knew (in spite of Plato's implied attribution of the discovery to Anaxagoras: *Cratylus* 409a–b; cf. 59 B 18 DK) that the moon shone by reflected light and that its phases were due to its orientation relative to the sun, and hence that it was spherical (cf. Empedocles, 31 B 41–7 DK), difficult though this is to reconcile with the
end p.37

Anaximandrean idea of the heavenly bodies as ‘garlands’, however that is to be construed.

(e) Later Eleatics

Parmenides sought to prove, on the basis of reason alone, that whatever existed could not change. His most celebrated follower, Zeno (born *c*.490 bc), attempted to buttress Parmenides' position by elaborating a series of paradoxes designed to demonstrate the incoherence of the concepts of motion and plurality. The former are well known, and require no rehearsal here. In the latter, Zeno argues that if a finite magnitude *m* is divisible, it must be so into either (*a*) finitely many or (*b*) infinitely many parts. But if (*a*), then what remains after the division will have a determinate magnitude, and hence ought to be divisible; but if (*b*), then the minimal parts must themselves either have some magnitude (in which case their sum will be infinite, contrary to the assumption that *m* is finite), or have no magnitude (in which case they sum to nothing at all). Any physicist wishing to rehabilitate the basic notions that underwrite our ability to describe the world around us must confront antinomies such as these (as well as the motion paradoxes), either by accepting, with (*a*) that some magnitudes are indivisible (as Democritus and the Atomists were to do: VI.1), or by way of a more subtle analysis of the concept of infinite divisibility which will defuse the paradoxical results of (*b*).

Melissus of Samos (fl. *c*.440 bc) was also a Parmenidean:

43what is always was what it was and always will be. For if it had come into being, necessarily before coming into being it would have been nothing. Yet if it had been nothing in no way would anything have come to be from nothing. (Simplicius, *On the ‘Physics’* 1. 3. 162. 24–6 = 30 B 1 DK = 525 KRS)

That is familiar stuff; but a later fragment introduces novel considerations:

44nor can it change in arrangement. For the previously existing arrangement (*kosmos*) is not destroyed, nor does a non-existent arrangement come into being. (ibid. 111. 24–6 = 30 B 7 (3) DK = 533 KRS)

44 specifically denies that new structures can emerge upon the rearrangement of pre-existing material elements, and that change and ultimately generation are explicable in terms of material rearrangement. His argument rests upon CP1*b* and its corollary

(CP2) nothing can be completely annihilated;

but Melissus further assumes, unwarrantedly, that CP1 and CP2 apply not only to material contexts but also to formal or structural domains. An

arrangement is simply a certain ordering of components: it is not the case that there are two objects, the set of components and the form, which are somehow glued together. Rather, form supervenes upon, or emerges from, the material components. But the concept of emergence is tricky, and requires careful handling; and subsequent pluralists would have to wrestle with its complexities.[50]

Finally, Melissus denied the possibility of motion on the grounds that it presupposes the existence of empty space or void: but void is what is not, and what is not cannot exist; hence there can be no motion (30 B 7 DK = 534 KRS). This argument was to prove immensely philosophically fertile: subsequent champions of motion either had to contend (with Empedocles, Aristotle, and Strato: §4*c* below; V.2*b*, 3*c*) that it did not, after all, presuppose void; or alternatively attempt to show (with Leucippus, Democritus, and Epicurus: VI.1*a*, 2*c*) that the notion of void was not, after all, incoherent.

4. Pluralism and Emergence

(a) Pythagoreanism

Pythagoras was already a legendary figure in Plato's time. He was born in Samos around 570 bc, from where he decamped to Croton some thirty years later. He was a complex mixture of mathematician and mystic; he believed in the transmigration of souls, and was for this reason a vegetarian. He also believed the soul to be immortal, although as far as we can tell his reasons for so doing were, if anything, empirical (he claimed to remember his past lives: Diogenes, *Lives* 8. 5). On the other hand, later Pythagoreans set much store by mathematics, in particular the study of ratios:

45 the Pythagoreans like to say that ‘everything is assimilated to number’, and sometimes to swear this most fundamental oath ‘by him who gave to us the tetractys which holds the source and root of ever-flowing nature.’ . . . The ‘tetractys’ is a number composed of the four primary numbers to produce the perfect number, namely ten: for one and two and three and four make ten. This number is the first tetractys and is called ‘the source of everflowing nature’ on the grounds that the entire universe is arranged harmoniously, while a harmony is a system of three concords, the fourth, the fifth and the octave; and the ratios of these three concords are found in the four numbers just mentioned, in one, two, three, and four. (Sextus Empiricus, *Against the Professors* 7. 94–5 = 279 KRS)

end p.39

The ratio of a string (or pipe) of a given length and that of one one octave higher is 2 : 1; similarly the ratio of the fifth is 3 : 2, and the fourth 4 : 3. This startling discovery of number in music suggested that mathematical relations might underlie physical reality in hitherto unsuspected ways, although much of their numerology was fanciful in the extreme (three is the number of the male, two the female: hence five is the number of marriage).

Aristotle ridiculed the Pythagorean notion that things really *were* numbers as an absurd category-mistake (*Metaphysics* 13. 8. 1083^{b}8–19 = 432 KRS). Yet he also wrote less dismissively:

46 since of these things [i.e. principles] numbers are primary by nature, and in numbers they thought they saw many resemblances to the things that exist and are generated, more so than in fire and earth and water . . . and since they saw that the attributes and the ratios of the attunements were numerical, and since everything else seemed to assimilate its whole nature to numbers, they supposed that the elements of numbers were the elements of everything, and the whole universe to be an attunement and a number. . . . Clearly then these thinkers also suppose that number is the principle both as matter for all things and as forming their attributes and permanent states, the elements of number being the odd and the even, the one limited, the other unlimited, and the one formed of both of them. (ibid. 1. 5. $985^{b}26$–$986^{a}20$ = 58 B 4, 5 DK = 430 KRS).

These views are usually associated with the Pythagoreanism of Philolaus of Croton (fl. *c.*430 bc).[51] In a fragment now generally accepted as genuine, Philolaus holds that

47 everything that exists must either all be limiting or unlimited or both limiting and unlimited. But they cannot only be unlimited. Now since it is evident that existing things come neither solely from limiting things nor from unlimited things, it is thus evident that the world and its contents were put together from both limiting and unlimited things (*apeira*). (Stobaeus, *Anthology* 1. 21. 7a = 44 B 2 DK = 425 KRS)

The 'limiters and the unlimited' are elsewhere connected with numbers, while the last sentence of **46** shows that the Pythagoreans assimilated the 'unlimited' to the even and the 'limited' to the odd (cf. Plato, *Philebus* 16c–18b: III.4*d*; Speusippus and Xenocrates, X.1*a*). Philolaus clearly sees number theory as having metaphysical import:

48 it was not possible for any of the things that exist and are known by us to have come into being without the existence of the essences of the things from which the cosmos was put together, both the limiters and the unlimited. And since these principles (*archai*) existed being neither similar nor of the same kind, it

end p.40

would have been impossible for them to have been put into order unless attunement had supervened upon them. (Stobaeus, *Anthology* 1. 21. 7d = 44 B 6 DK = 429 KRS)

Whatever the obscurities of that, Philolaus is clearly trying to answer to one of the constraints generated by the Parmenidean critique, namely that of no genuine creation *ex nihilo*, or CP1. And however confused and unclear their doctrines, the Pythagoreans begin to grasp the importance of the notion of form and structure, as well as the possibility of applying mathematics to the explanation of the physical world. Pythagorean numerology led to a novel astronomical postulate:

49 while most say that the earth is in the centre . . . the Italian philosophers known as Pythagoreans say the opposite; for they hold that fire occupies the centre, while the earth, being one of the stars, makes night and day by being carried round the centre in a circle. Moreover they construct another earth opposite this one which they call 'counter-earth', not seeking accounts and explanations in relation to the phenomena, but rather trying to drag the phenomena into conformity with certain accounts and beliefs of their own. And there are others to whom it does not seem right to assign the central position to the earth, seeking justification not from the phenomena, but rather from their arguments. (Aristotle, *On the Heavens* 2. 13. $293^{a}18$–30 = 58 B 37 DK 446 KRS)

The bizarre theory of the counter-earth may well be Philolaean (Aëtius 2. 7. 7; cf. I.4*a*). But while Aristotle's suggestion that its postulation was motivated by a priori considerations is no doubt justified (it brings the heavenly bodies up to the perfect number ten), and while it does contrast with Aristotle's own insistence that in such matters explanations should be tentative and driven by the phenomena (*Meteorology* 1. 7. $344^{a}5$–8), none the less the counter-earth may also have been designed to fulfil the same role as Anaxagoras' invisible bodies (**50** below) in explaining the relative frequency of lunar eclipses (cf. *On the Heavens* 2. 13. $293^{b}21$–5).

Even so, it is difficult to see how the account could have been worked out in detail. The earth revolves around a central fire (the sun is a mirror reflecting its light: Aëtius 2. 20. 12), which is never directly visible since we live on the face pointing outwards from the centre. The counter-earth is supposedly invisible either because it is located between the earth and the central fire, or because it orbits the fire diametrically opposite the earth (Aëtius 3. 11. 3); and in either case, on any account of the relative orbits of the sun, moon, and earth, it is hard to see how the counter-earth could interpose itself between sun and moon; and this is not the only difficulty with the theory.[52]

end p.41

Like so much else in Pythagoreanism, in the case of Philolaean astronomy it is hard to see where honest scientific speculation ends and mysticism begins. That the theory was intended in some degree to account for the celestial appearances is fairly obvious; but it is extremely inadequate, and is forced to ignore several astronomical facts which were certainly known at the time (notably the inclination of the ecliptic). It is hard, in the end, to quarrel with Aristotle's tart judgement.

(b) Anaxagoras: Mixture and Arrangement

Anaxagoras (*c.*500–428) was also interested in astronomical and cosmological matters (he held that the sun was a hot stone, and was prosecuted for impiety for his pains: Diogenes, *Lives* 2. 12; cf. Plato, *Apology* 26d):

50 he says that the sun and the moon and all the stars are fiery stones, which are carried around along with the rotation of the ether, and that there are some bodies below the stars which are carried round by the sun and the moon, although they are invisible to us. . . . The moon is lower than the earth and nearer to us. The sun exceeds the Peloponnese in size. The moon's light is not its own, but comes from the sun. The moon is eclipsed when it is screened by the earth, and sometimes by the bodies beneath the moon; the sun when it is screened by the new moon. (Hippolytus, *Refutation* 1. 8. 6–9 = 502 KRS)

That passage shows Anaxagoras to be concerned with the explanation of astronomical phenomena, notably eclipses, and to have proceeded well on the way to developing the correct account of them. The mysterious invisible bodies also function in the explanation of eclipse phenomena: Anaxagoras presumably knew that lunar eclipses occur with greater frequency than solar ones, and thought that he needed as a result to postulate other bodies which might on occasion occult the moon. Diogenes (*Lives* 2. 10) records the story that Anaxagoras had predicted the fall of a meteorite: and while evidently apocryphal, the tale suggests his close association with astronomical and cosmological inquiry, an association underlined by Plato's critique of his method in the *Phaedo* (97b–99a: III.1*a*, **122–3**).

In physics, Anaxagoras was painted by the ancients as an Anaximenean; yet air is not primary for him, nor does he take all material objects to be merely modifications of it. Indeed, his physics is strikingly original:

51 everything was together, infinite in both number and smallness; for the small too was infinite; and while everything was together, none of them were evident because of their smallness. For *aēr* and *aithēr* controlled all things, both being infinite. (Simplicius, *On the 'Physics'* 1. 4. 155. 26–9 = 59 B 1 DK = 467 KRS)

end p.42

52 But before these things [i.e. ordinary stuffs and objects] were separated off, everything was similar, and there was no evident colour: for the mixture of everything prevented it, the wet and the dry, the hot and the cold, the bright and the dark, since there was much earth in the mixture, and innumerable seeds in no way like one another. (ibid. 1. 2. 34. 21–5 = 59 B 4 DK = 468 KRS)

Several fragments assert that 'in everything there is everything' (59 B 6, 11 DK, etc.), and not just at the beginning of the cosmogony but at all times. This is probably to be interpreted as the radical thesis that in any portion of stuff there are traces of all other (basic) stuffs.[53]

There is, however, one exception to this universal intermixture (hence-forth UI):

53 the other things have a share of everything, but Mind is infinite and autonomous and mixed with no other thing. For if it were not by itself but mixed with something else, it would have a share of all things if it had a share of any, since in everything there is a share of everything . . . and the things that were mixed with it would hinder it so that it could not control anything in the way it does now, being alone by itself. (ibid. 1. 4. 164. 24, 156. 13–19 = 59 B 12 DK = 476 KRS)

Later philosophers did not think much of Anaxagoras' Mind: both Plato (III.1*a*, **122–3**), and Aristotle (*Metaphysics* 1. 4. 985^{a}18–22: IV.3*a*) dismiss it as explanatorily inadequate. But Aristotle commends Anaxagoras, albeit somewhat backhandedly (he says that he resembles a sober man among a crowd of drunks: ibid. 1. 3. 984^{b}14), for having introduced a directing force into cosmology. Aristotle criticized the early Presocratics' essays in physical explanation for concentrating solely on material causes and explanations; at least Anaxagoras perceived the necessity of injecting some ordering force into the cosmos. It is not enough to *say* that things are caused to move by Anaximandrean eternal motions, or the like. The emergence of order and structure in the cosmos requires explanation; Mind is responsible for separating things off from the original smooth mixture in a series of vortices.

But in what sense does everything else contain a portion of everything? Anaxagoras is no atomist: things are infinitely divisible, and at each stage of the division what remains is more of the same stuff to be divided. Stuffs are uniform (in Aristotle's jargon, homoeomerous). A stuff *S* is uniform just in case any part *P* of *S* is also *S*. In contemporary physics, mass-terms are restrictedly uniform (any macroscopic part of the water in a glass is water) but that uniformity breaks down at the molecular level. Crucially,

end p.43

for Anaxagoras, it does not. There are no minimal parts of basic, unadulterated stuff. This Principle of Uniformity (PU) threatens the coherence of the notion of basic stuffs: if they can never be isolated, in what sense can they be said to exist? Moreover, UI and PU seem incompatible. This lump of chalk, by UI, contains cheese; furthermore, by PU, every piece of it contains cheese. But the contained cheese itself contains chalk, and portions of everything else. The difficulty is occasioned by thinking of the contained cheese as being in the form of microscopic lumps of pure gorgonzola (if such a thing is possible: blue cheeses in particular present obvious problems for the notion of homoeomeriety); rather cheese is present in a certain proportion in the smooth mixture that is ordinary chalk at any and all levels of dissection.[54]

However, it does emerge that what we *call* chalk (i.e. the stick in my hand) is really only *predominantly* chalk: there is more chalk in it than there is any other elemental stuff. 'Chalk' is thus ambiguous; what we mean (ordinarily) by chalk is not elemental chalk. That is, I think, harmless enough, although it makes Anaxagoras' elemental stuffs theoretical postulates. Anaxagoras finds himself forced to adopt such postulates as a result of accepting a Parmenidean regulative principle, namely CP1*b* (and its corollary CP2):

54that there is generation and destruction is not rightly thought by the Greeks, since no thing either comes to be or is destroyed, but is commingled and separated off from existing things: for this reason one may rightly call generation commingling and destruction separation. (Simplicius, *On the 'Physics'* 1. 4. 163. 20–4 = 59 B 17 DK = 469 KRS)

Strictly speaking, there is no generation and destruction, only rearrangement; and the proper lesson to be drawn from Parmenides is the universality of the Principle of Conservation. Thus I derive chalk from cheese by concentrating the cheese and separating off the chalk. More pertinently, bread becomes flesh because the flesh contained in it is skimmed out. Thus new stuffs (and their properties) emerge only in a weak (if literal) sense; there are, properly speaking, no new stuffs and properties at all (59 B 10 DK). Anaxagoras' physics also allows us to demystify the emergence of quality, since UI applies to qualities too: Anaxagoras explicitly includes colours as being among the things which are in everything (59 B 10, 21 DK). In a famous epistemological fragment, Anaxagoras declares that

end p.44

'what appears is a glimpse of what is unseen' (59 B 21a DK: VI.1*a*, **214**): and he means that in a very real sense.[55]

But what about emergence of structure of the type stigmatized by Melissus in **44**? How are we to account for the appearance of form? **44** is usually taken as simply extending the Eleatic rejection of generation from non-being to structures and arrangements. Yet it may equally embody the well-motivated demand that such emergence should be causally explicable: but again, given PCS, it is hard to see how it can be, unless the forms in question somehow precede their particular instantiations. This is no trifling problem.

Both Plato and Aristotle offered (strikingly different) answers to it, and modern genetic theory may be seen as a response to one particular case of it. It is not enough, with Anaximander, simply to treat movement as a fundamental, irreducible property; we must account too for the extraordinary regularity and repetitiousness of the world's processes. This explanatory requirement lies behind one of the crucial debates in the history of explanation, that between mechanism and teleology. Can we explain the world and its workings solely on the basis of the mechanical interaction of its fundamental constituents? Or must we invoke some further basic principle of goal-directedness in order to account for its structural stability? Aristotle and Plato both (although again differently) thought the latter, which is why they applaud Anaxagoras for making Mind basic, even while finding his physics excessively mechanistic in tenor.

One feature of that leaning towards mechanism was his doctrine of seeds (**52**). These, apparently, include any microscopic template for the generation of macroscopic objects, and thus include, but are not exhausted by, seeds in the usual biological sense.[56] Anaxagoras' account of structure is thus basically preformationist, although its details are obscure, and it contains obvious difficulties: surely not *every* structure, including that of artefacts, can be so preformed. But whatever is to be said about that, Anaxagoras' physics represents a sustained attempt to build an account of a mobile, multifarious world compatible with Parmenides' logical and metaphysical constraints.

(c) Empedocles: Structure and Emergence

Empedocles (*c.*495–*c.*435) also felt the need to accommodate Parmenides' metaphysical insights; and he too saw the need for motive forces in his

cosmos. Indeed, he postulated two of them: Love, which tends to bring things to maximum unity and cohesiveness, and Strife, which breaks them down again. The two wrestle forever in the cosmic ring: sometimes Love triumphs completely, forming the universe into a single, homogeneous sphere, only for Strife to begin breaking everything down again (31 B 17, 21, 28, 29 DK). What persists are the forces, and the four elements, or 'roots' as Empedocles called them (31 B 6 DK), earth, water, air, and fire (31 B17, 22, 98 DK); and this persistence is the proper lesson to be drawn from Eleatic logic:

55 birth has none of all
The mortal things, nor any end of accursed death,
But only mixture and interchange of what is mixed
Exists; birth is the name given to them by men.
(Plutarch, *Against Colotes* 1111f = 31 B 8 DK = 350 KRS; cf. 31 B 15 DK)

Indeed

56 From the non-existent it is impossible to come to be
And that what is be destroyed cannot be done or countenanced;
For wherever one may put it, there it will always be.
(ps.-Aristotle, *On Melissus, Xenophanes, and Gorgias* 975^b1–3 = 31 B 12 DK = 353 KRS)

Empedocles' language unmistakably echoes that of Parmenides. Motion, and hence change, are to be rehabilitated against the Melissan attack (§3*e* above) by denying that void is a necessary condition for motion: things can move by serial replacement, or *antiperistasis* (ibid. $976^{b}22$–9; cf. Aëtius 2. 7. 6); this solution was adopted and buttressed by Aristotle and Strato (V.2*b*).[57] The elements themselves, in various proportions, combine to produce the ordinary stuffs of the world:

57 And gracious earth received in her broad hollows
Two parts of the eight of the gleam of Nestis [i.e. water].
Four of Hephaestus [i.e. fire]; and white bones came to be
Wondrously joined by the bonds of harmony.
(Simplicius, *On the 'Physics'* 2. 3. 300. 21–4 = 31 B 96 DK = 374 KRS)

58 And earth happened in roughly equal quantity on these,
Hephaestus, rain, and all-flashing *aithēr*
Anchored in the perfect harbours of Love,
Either a little more or less in more of them,
And from them came blood and the other forms of flesh.
(1. 2. 32. 6–10 = 31 B 98 DK = 373 KRS)

end p.46

Shorn of their poetic coat, **57** and **58** offer a combinatorial theory of chemical composition. Bone is two parts earth, two parts water, four parts fire, held together by Love.[58] Blood is a more or less homogeneous mixture of all the elements, slight variations accounting for its slightly different forms. Empedocles seeks to marry traditional elemental physics with a Pythagorean predilection for mathematical structure. This compositional physics accounts not merely for the generation of compound stuffs, but also, ultimately, for form:

59 In Anger [i.e. Strife] all are of different forms and separate,
But in Love they [i.e. the elements] come together in mutual desire.
From them come everything that was and is and will be;
Trees have sprouted, and men and women,
Beasts and birds and water-nourished fish,
And the long-lived gods, greatest in honour.
For there are only these things [elements], but running through one another
They become other-faced—so much does mixture alter them.
(1. 4. 159. 19–26 = 31 B 21 DK = 355 KRS)

At bottom there are just the elements and their arrangements: but together they are responsible for the bewildering variety of form the world exhibits. Elsewhere, he employs an illustrative analogy:

60 As when painters skilfully craft votive tablets . . .

When they grasp many-coloured pigments in their hands
And mix them with harmony, some more, some less,
They create from them forms alike to everything,
Crafting trees and men and women.
(160. 1–6, 31 B 23 DK = 356 KRS)

Paints mixed on a palette may produce entirely new shades; and those shades may then be applied to the canvas in the likenesses of things. **60** clearly adverts to the generation of both quality and structure. But how do the elements become 'other-faced'?

To merit the description 'emergentist' Empedocles would have to agree (with some modern philosophers of biology) that the new properties are not in any real sense reducible to those of the basic elements from which they emerge: what occurs is a genuine ontological novelty, one which can be predicted and placed in lawlike correspondences, but which resists metaphysical reduction. But this apparently contradicts CP1*b*, which Empedocles explicitly endorses (**56** above), and also seems inconsistent

with his remarks to the effect 'nothing comes to be over and above these things [i.e. the basic components of his ontology]' (31 B 17 DK).

We need briefly to examine another aspect of his physical theory. Empedocles posited the existence of passages (*poroi*) in things, of varying gauges, to account (among other things) for perception, magnetism, growth, respiration, and, pre-eminently, physical mixture. The doctrine is dark, and the evidence largely derived at second hand from Aristotle's *On Generation and Corruption* 1. 8, and Theophrastus' *On the Senses* 7–24; but he apparently envisaged the material elements as able to interlace (the metaphor is Mourelatos's) with one another in such a way as to produce new substances. Not for Empedocles, then, Anaxagorean complete inter-penetration: rather the water remains in the bone *as water*; and we may tentatively ascribe its Empedoclean adhesive properties (n. 57) to the particular structure of its lattice.

There remain the problems of explaining how these new compounds display apparently new properties. Empedocles might have adopted Democritus' approach (VI.1*a*, **215**), simply denying colours any genuine reality. But that Democritean solution, fine in principle if applied to all colours as secondary properties, is less satisfactory when applied to only some of them (essentially the colours of mixture). Empedocles may have thought mixed colours illusory, in the way that the *pointilliste*'s green is an illusion caused by the contiguity of tiny yellow and blue dots; but it is hard to generalize this account to cover other physical properties of compounds (flexibility, ductility, and so on).[59]

That leaves form; and raises the issue of how Empedocles sought to account for the regularity of nature, particularly in the biosphere. Aristotle was to chide him, along with Anaxagoras and Democritus (cf. VI.1*d*), for refusing to acknowledge the world's teleological aspect, which was central to Aristotle's own picture (see IV.1*a*, 3*a*–*e*). In particular, he castigates Empedocles for leaving things to chance. Simplicius (*On the 'Physics'* 2. 4. 330. 30–331. 14) reports a number of Empedoclean examples of things ascribed to chance, in particular animal generation: and he concludes, dismissively:

‘Empedocles, who seems to use chance only in small matters, is not worthy of much attention, not having explained what chance is’ (331. 15–16).

This is not entirely fair. Empedocles outlined a remarkable theory of the evolution of animals, in which animal parts were first formed directly from the elements, then became joined more or less haphazardly, among which some stable ‘whole-natured forms’ arose from them which were able to reproduce sexually (Aëtius 5. 19. 5 = 31 A 72 DK; Simplicius (*On ‘On the*

end p.48

Heavens’ 586. 6–87. 26: 31 B 35, 57, 59 DK; cf. 31 B 60–2). The generation is thus initially random, but some of the generated beings prove fit to survive and reproduce. It appears that Empedocles had glimpsed, however dimly, the crucial fact that certain chance arrangements of things can turn out to be self-replicating. Their original generation was, in the appropriate sense, chance: but their continued success is to be explained in terms of the mechanical properties of the successful combinations. Empedocles' theory has some claim to be a distant ancestor of the enormously powerful and successful explanatory paradigm of natural selection.

(d) Coda: The Atomists

The other great late Presocratic attempt to evade the grasp of Eleatic argument is that of the original Atomists, Leucippus and Democritus. For strategic reasons of exposition, I treat their views more fully in VI.1; but a brief outline of their position will not be out of place here.

Anaxagoras and Empedocles try, in their very different ways, to rebut Eleaticism from within the confines of a continuous physics. This is evidently the case with Anaxagoras; but it is equally true of Empedocles: their worlds contain no gaps, and their stuffs are infinitely dissectible. By contrast the Atomists' universe is discontinuous: there are minimal bits of stuff (or rather of distinct stuffs), the atoms, which cannot be further subdivided, and are indeed impervious to change of any kind (the atoms are in fact minimal Parmenidean objects: VI.1*c*). Moreover, space is itself composed of minimal parts, corresponding to the minimal parts of the atoms themselves; and for this reason, Zenonian paradoxes of motion and plurality are both blocked.

Equally, the Atomists accept the Eleatic notion that motion is impossible without a void, but they employ this contention to argue for the real existence of void (VI.1*a*). Generation and decay are simply the arrangement, rearrangement, and dissolution of the individual unchanging atoms into macroscopic bodies; thus the Atomists, like Anaxagoras and Empedocles, seek to retain a sense in which the Eleatic proscription on change and generation turns out true (VI.1*d*).

Finally, the atoms are infinite in number, as are the worlds which are generated from them. These worlds arise both by chance and by necessity, by chance in the sense that no final causation or immanent tendency to form impels them to come into being (compare the views of Plato and Aristotle: III.4*a*; IV.1*a*, 3*a*–*e*); by necessity in the sense that their

coming into being is the necessary outcome of particular, albeit random, atomic interactions (VI.1*e*); the mechanism of the Atomists requires no directing force, or Anaxagorean Mind.
end p.49

5. Conclusions

Aristotle thought that Thales and his successors had wrought a revolution in thought: and he was right. If their attempts at physical explanation were naïve, and excessively concerned with the material at the expense of other causal factors (see **1** and **2** above; and cf. IV.2*a*), at least they *were* attempts at physical explanation, and ones which demonstrated a steadily increasing sophistication from Thales through Anaximander and Anaximenes (§§2*b–i*). If Xenophanes sounded a cautionary note of epistemological scepticism, it did not prevent him from pursuing his own projects, in both cosmology and theology (§3*a*). Heraclitus puzzled more deeply than any of his predecessors over the notion of change, sketched a rudimentary metaphysics of perpetual flux, and expanded on the Anaximandrean theme of essential balance in the universe, while also occupying himself with the analysis of the human soul (§3*b*), a concern also evinced by Alcmaeon (§3*c*).

The general confidence of the early Presocratics in the possibility of their natural investigations was rudely shattered by Parmenides' arguments establishing the conceptual impossibility of change and generation on the basis (among other things) of certain intuitively powerful a priori theses regarding causation (§3*d*), arguments buttressed by his followers Zeno and Melissus (§3*e*). In the period of retrenchment and reassessment which follows, Anaxagoras, Empedocles, and the Atomists all seek to find ways to evade the unpalatable conclusions of Eleatic argument, while accepting and adapting the theoretical constraints upon the notions of causation and generation which underpinned it (§§4*b–d*; cf. VI.1). The result is physical theories which, for all their mutual differences (not to mention their shortcomings), are none the less united by their much more sophisticated and theoretically dense conceptual structures than anything which has gone before.
end p.50

II Science and Sophistry

R. J. Hankinson

1. Rational Medicine

(a) Naturalistic Explanation

Medicine is practised in every society. In the Homeric poems we find the healers Machaon and Podalirius; Greek myth speaks of Chiron the surgical centaur; and Greek doctors liked to trace their ancestry back to the divine founder and patron of the art of medicine, the god Asclepius. Yet a remarkable development begins to take place in medical practice towards the end of the fifth century bc . The old methods and prescriptions of the temple doctors (the wearing of magic amulets, divination on the basis of dreams experienced while sleeping in sacred sanctuaries, prayer) were challenged by a new, rational medical paradigm. Medicine has already played a role in philosophy; but it is in the group of treatises associated with the semi-legendary figure of Hippocrates[1] that the revolutionary aspects of the new medicine are most clearly visible.

The archaic notion of disease parallels early accounts of other physical phenomena; typical is the celebrated description of the pestilence visited on the Greek army at Troy by an angry Apollo (*Iliad* 1–187). That may be contrasted with Thucydides' description of the great Athenian plague of 430 bc in *The Peloponnesian War* (2. 47–54). It began in Ethiopia, spreading through Libya and Egypt into the Persian empire, but

61as to the question of how it could have first come about or what causes can be found adequate to explain its powerful effect on nature, I must leave that to be treated by other writers. . . . I shall simply describe what it was like and set down the symptoms, knowledge of which will enable it to be recognized should it ever break out again. (2. 48)

Medicine was powerless (doctors in particular suffered through their contact with the sick); but

62nor was any other human art or science of any use at all. Equally useless were prayers offered in the temples, consultation of oracles, and so on; indeed, in the

end p.51

end people were so overcome by their sufferings that they forgot about such things. (2. 47)

Thucydides then describes in detail the physical signs and symptoms associated with the epidemic. In general it was not a bad year for disease,

63but those who did have any prior illness all eventually caught the plague; but in other cases there seemed to be no occasion (*prophasis*) for the attacks. (2. 49)

He gives a harrowing account of the progress of the disease (2. 49–51), although it was not uniformly fatal (Thucydides himself survived an attack), and those who recovered developed an immunity to it. He further documents the plague's social effects:

64no fear of gods or law of man had any restraining influence; as for the gods, it seemed to be the same thing whether one worshipped them or not, when one saw the good and the bad dying indiscriminately. (2. 53)

Traditional religious remedies and practices were seen to be useless, and abandoned: evident inefficacy induces scepticism.[2]

Thucydides' treatment of the plague is entirely empirical in nature. He does not speculate on the plague's causes (**62**), beyond saying that it frequently struck with no obvious antecedent occasion (**63**). Yet while the majority of the rationalist doctors of the age were concerned with the causal explanation of disease, seeking to uncover the deep physiological and pathological structures in virtue of which people were healthy and sick, medicine is a field in which empirical results count. No practising doctor, however theoretically sophisticated, can afford to ignore the actual phenomena that manifest themselves in particular cases. Thus medicine, a practical discipline, but one which was developing theoretical pretensions to understanding, is the natural arena in which the competing demands of empirical adequacy (both in terms of prognosis and cure) on the one hand and of theoretical depth and enlightenment on the other are most clearly to play themselves out, in the course of a long and rich history of interplay between speculative philosophy and pragmatic medicine (see further Chapters IX, XI).

Although himself no doctor, Thucydides' detached, sceptical approach is typical of the new attitude in medicine, which is also clearly exemplified in the roughly contemporary (*c*.420 bc) Hippocratic treatise *The Sacred Disease*:

65 I intend to discuss the so-called ‘sacred disease’; it does not seem to me that it is any more divine or sacred than any other disease, but it has a nature (*phusis*) and an occasion (*prophasis*), but men suppose it to be a divine thing because of its strangeness and their helplessness in the face of it. (*The Sacred Disease* 1. 1–5 Jones)[3]

The ‘sacred disease’ is epilepsy (and other types of seizure), traditionally thought to be visited by the gods. Our author thinks it no more divine than malarial fevers (1. 10–18; cf. 5. 1–2, 21. 4–8), although he does not deny its divinity: all diseases are (in a sense) ‘both divine and human’ (21. 8). But they all have a *phusis*, a nature or internal constitution, and they arise from identifiable origins:

66 each thing has its own nature and power (*dunamis*); and none are unintelligible or beyond therapy. And most are cured by the same things that cause them. (21. 8–12)

Hippocratic medicine is strongly allopathic (the slogan ‘opposites cure opposites’ is ubiquitous throughout the corpus), and this is, first homeopathic appearances notwithstanding, the sense of that second sentence: once you have identified the pathogen, a cure can be effected by removing or suppressing it.

The allopathic principle is intended to regulate treatment; but it is related to the general causal principle

(CP3) For any *x*, any *y*, and any *F*, if *x* is the cause of *y*'s being *F*, then removal of *x* will contribute to the suppression of *F*.

CP3 does not claim that removal of the pathogenic factor will effect a cure by itself, since the disease may have established itself in the body independently of its original cause, and thus require some active antidote. In that case, allopathy will dictate that, if the disease consists in an imbalance of the body's basic constituents in a particular direction, it must be corrected by applying the opposite substance or quality to the one which now predominates. The paradigm of disease as imbalance (which perhaps originated with Alcmaeon: I.3*c*), and the associated allopathic principle of treatment, became dominant (although not unchallenged) throughout the history of Greek medicine. Both the structure of the pathological theory, and the insistence upon the intelligibility of disease and bodily function and dysfunction, are products of the same intellectual attitudes as those which nourished Presocratic speculation.

The ‘sacred disease’ is thus both naturally explicable and treatable: those who label it divine are charlatans, seeking to excuse their own ineptitude. Moreover, their prescriptions are mumbo-jumbo (2. 1–32). The
end p.53

‘magicians’ ban goat products, in response to which the author remarks sardonically:

67 I suppose no Libyan dwelling in the interior can enjoy good health, since they lie on goatskins and eat goat meat, and have no blanket, cloak, or shoe that is not made of goat. . . . But even if eating or using these things were to nourish and increase the disease, while not eating them cured it, the god would still not be responsible (*aitios*) for it, nor would the purifications help; rather the foods would be both curative and harmful, and the alleged divine power disappears altogether. (2. 35–46)

The first sentence demonstrates a concern with explanatory generality, and implicitly invokes the regularity of causes. If goat products really are relevant to epilepsy, there should be more of it in Libya. This argument is inconclusive: the Libyans might avoid a higher incidence of the disease for other reasons (they may consume antidotes, or have a natural immunity; or perhaps they simply propitiate the appropriate gods better). But such empirical considerations are still relevant, and ignored by the champions of the supernatural model.

But even if goat products were harmful, that fact would indicate that the disease had natural rather than supernatural causes. Nor can the sacred theory explain the hereditary nature of the disease, or why it particularly attacks people of phlegmatic (i.e. cold and moist) constitution:

68 if it were more divine than the others, it should attack all indiscriminately, rather than differentiating between bilious and phlegmatic. (5. 17–21)

Epilepsy is in fact (he thinks) a disease of the brain (5. 17–20), whose cerebral functions are impaired if the flow of air to it is blocked, as can be caused by an excess of phlegm (5. 6–10); and he tries (admittedly in fantastical fashion) to explain why its incidence is different among different age groups, and different climates (5. 11–16).

Sacred Disease recalls Presocratic naturalism in its pretensions to explain (and indeed control) physical phenomena on the basis of an abstract and general causal theory, yet one which seems none the less wholly to outrun its evidential base. Although it is loosely based upon observations, and while it is designed to account for certain facts which the sacred theory cannot (such as its heritability, and the fact that it affects young people to a disproportionate extent), the profane, rational account is itself mumbo- jumbo. The avowed method of seeking causally relevant conjunctions of general facts is poorly prosecuted in practice, and in many cases it is unclear if the alleged phenomena upon which the theory is based could even have been observed, much less that they support the author's own physical hypotheses at the expense of any others.
end p.54

(b) Causes and Constitutions

The writer of the *Sacred Disease* distinguishes between the internal structure of a disease and what occasions it (**65**, **66**); Thucydides noted that no originating causes were associated with the plague, although certain conditions rendered people more prone to it (**63**). The Hippocratic writers are the first explicitly to distinguish between those internal, constitutional factors which render some people more susceptible to particular diseases than others, and the external triggering causes which set the pathogenic process in motion. This coincides with the emergence of the goal of explanatory generality: the rationalist wants to know why it is that diseases affect some people and not others, and why some recover but others do not. Such concerns are practical, but they also involve an abstract desire to render the world intelligible:

69 I see that doctors are not experienced in the proper way to distinguish the various weaknesses that occur in diseases, those caused by starvation, or some other provocation, or pain, or by the acuteness of the disease; nor can they distinguish the various affections in all their forms which our natures and habits engender in each of us. Yet knowledge and ignorance of such things brings life or death. (*Regimen in Acute Diseases* 43. 1–8 Jones)

The bulk of *Regimen in Acute Diseases* consists of detailed and precise symptomologies for the various different (and, as the author thinks, distinct) types of serious illness, and reports of the effects of different prescribed remedies and foods. The remark about natures and habits is important. It is an article of Hippocratic faith that individuals differ from one another both in the basic constitutions of their bodies and in the extent to which their bodies have become habituated by regimen to a particular lifestyle. If people have grown accustomed to eating two meals a day, they suffer seriously if they miss lunch, as the more abstemious do not (43. 30); and bathing is indicated in most acute diseases provided that the patient is used to it (43. 65–8). In general

70 this is a sufficient sign that the greatest changes in our natures and habits are the most serious causes of disease. (35. 1–4)

The author wishes to defend medicine against its detractors (cf. §2*c* below), whose charges have been lent substance by incompetence and disagreement among doctors (35. 1–8) He is convinced that these disagreements can be resolved, and the proper therapies discovered, provided sufficient attention is paid to clinical detail. The emphasis on clinical detail, reminiscent of Thucydides' description of the plague, is important. *Regimen in Acute Diseases* has, unlike *The Sacred Disease*, no grand theory to offer; rather it contents itself with observation of what occurs in particular

types of cases. This concern with precise observation is best exemplified by *Epidemics* 1 and 3 (which are generally taken to comprise a single treatise).

The structure (if it can be called that) of *Epidemics* 1 and 3 is curious. The writer begins with two ‘constitutions’, or descriptions of general climatic conditions obtaining in a particular area (Thasos, an island in the northern Aegean) at particular times, and the peculiar types of sickness associated with them. The writer notes the prevailing winds, the dampness or dryness of the climate, and whether they contrast sharply with previously obtaining conditions. Next he catalogues the types of patient most afflicted: ‘the sufferers were youths, young men, and men in their prime, usually those who frequented the wrestling school and gymnasia; few women were attacked’ (*Epidemics* 1.

1. 20–2 Jones), before giving a detailed symptomology, carefully distinguishing different types of fever.

The second constitution follows the same pattern. Chronic, persistent complaints are differentiated from acute diseases which follow a definite pattern from onset through a period of worsening to the 'crisis' of the disease, which is followed by either recovery or death. Vomiting, sweating, diarrhoea, catarrh, urethritis, are particularly noted. Finally, the writer offers prognostic advice: the condition of the evacuated matter is the best predictor of the course of the disease (in general if it is 'raw', 'unconcocted', it bodes ill: the smoother and better metabolized the better). There is little attempt to indicate therapy:

> **71** say what has happened, diagnose the current condition, predict what will happen; take care of these things. In regard to diseases, practise two things, helping or at least not harming. The art (*technē*) has three parts to it: the disease, the patient, and the doctor. The doctor is the servant of the art; and the patient must help the doctor fight the disease. (ibid. 1. 11. 9–15)

This attitude is echoed in *On Prognosis*: the doctor gains a reputation for expertise by correct prediction, and comforts the patient and family thus. But many diseases are simply beyond medical help, and you will avoid censure if you say so frankly at the outset:

> **72** anyone who is going to make accurate forecasts as to who will recover and who will die, and whether the disease will last a greater or a lesser number of days, must understand how one can distinguish all the signs, estimate their relative powers, as I have described above, particularly in regard to urine and saliva. One must also immediately recognize the attacks of the endemic diseases, and not overlook the constitution of the season. Moreover, one must well understand the indications and other signs, that in every year and in every land the bad ones signify bad outcomes, the good good, since all the signs described above have manifestly the same significance in Libya, Delos, and Scythia. So one must well understand that there is nothing remarkable about being right in the same places in the vast preponderance of cases, if one learns how to

end p.56

> distinguish them and knows how to estimate them accurately. Do not ask for the names of any disease which I have left out of the account—for in all cases you will know by the same signs that the diseases will come to a crisis in the times mentioned. (*On Prognosis* 25. 1–24 Jones; cf. 1. 1–27)

Again, our author is concerned to stress the importance of the generality of the account. Even if local factors induce variations in the incidence of disease, some conditions are universal, and can be recognized as such on the basis of distinguishing signs. The writer offers no taxonomy of diseases as such, nor is he directly at least concerned with their physical explanations: but he is concerned to determine significant groups of symptoms.

The constitutions of *Epidemics* 1 are succeeded by fourteen individual case-studies, in each of which the patient's sex and domicile are recorded, followed by a selective account of the disease's progress and outcome. Sometimes a *prophasis*, or occasion for the disease, is suggested: 'after over-exertion, drinking, and intemperate exercise, he took a fever' (*Epidemics* 1. 26. ii); 'a man dined when over-heated and drank too much; during the night he vomited everything up; high fever, pain in right hypochondrium, with a soft internal inflammation; a disturbed night' (1. 26. xii). Such antecedent causes are also mentioned in some of the cases of *Epidemics* 3: 'Chaerion who lodged at Delias' house took a fever as a result of drinking' (*Epidemics* 3. 1. v); 'the lad lying sick by the Liars' Market took a fever as a result of exhaustion' (3. 1. viii); 'Pythion . . . had a violent rigor and high fever as a result of overwork, exhaustion, and insufficient attention to diet' (3. 17. iii); 'Delearces' wife . . . took a high fever as a result of grief' (3. 17. xv); and finally 'a young man who had been running a temperature for a long time as a result of drinking and sexual over- indulgence took to his bed. His symptoms were shivering, nausea, insomnia, and lack of thirst' (3. 17. xvi). Antecedent factors are not, however, mentioned in most cases: 'for some reason he took a fever and went to bed' (3. 1. iii). Moreover, our author writes:

73 in the early spring, just at the time when the cold snaps happened, there were many cases of serious erysipelas,[4] some with an originating cause (*prophasis*), some without. (3. 3. 1–3)

A *prophasis* is an *evident* cause, and a sign of what is likely to happen (cf. XI.2*b*). If the physician is aware of it, he will be able to modify his prognosis accordingly. Our authors offer no causal account of how the *prophaseis* operate; they simply register them and their apparent consequences, in a manner later adopted by the Empiricists (IX.2*a*–*b*, *e*). Sometimes too they note their apparent insignificance:

end p.57

74 In many of the cases where erysipelas occurred all over the body with an originating cause, it was a trivial accident, or a tiny wound, especially when the patients were around sixty years old, and the wound was in the head, even if it was only slightly neglected. (3. 4. 1–5)

A *prophasis* is, in ordinary Greek, frequently a pretext:

75 sometimes Hippocrates uses the term *prophasis* of those things that are falsely called causes, as is the general custom, but he often applied it to evident causes, and sometimes to any kind of cause in general. (Galen, *On Hippocrates' 'Epidemics'* 17a 52 Kühn = *CMG* v. 10. 1. 30.28 ff.)[5]

This 'general custom' is well attested (cf. e.g. Thucydides 3. 86 on the pretexts given by the Athenians for sending aid to Leontini), and is often explicitly contrasted with the real, ulterior motive (e.g. by Herodotus 1. 29 on Solon's real and pretended reasons for leaving Athens). But it need not have such connotations of falsity or fraudulence (cf. *Iliad* 19. 301–2, where the Greek women weep for their own losses, prompted by Patroclus' death, described as a *prophasis*); a *prophasis* is simply the ostensible reason or surface cause for something, as contrasted with its full cause or complete reason (generally denoted in the Hippocratic corpus by the term *aition*).[6]

Here in outline is a moderately complex aetiological theory. First there are the basic predispositions to various types of illness associated with different ages and sexes (the Hippocratics are particularly interested in gynaecology and obstetrics). These innate predispositions may be exacerbated by an over-indulgent lifestyle as well as by prevailing climatic conditions. Finally, any of these predispositions may be triggered into pathogenesis by some excess or deviation from the patient's normal regimen. The trigger events may be apparently negligible: what matters is that they can be seen to correspond with subsequent patterns of events. This model of health and disease was to remain, in its outlines at least, the dominant medical paradigm for more than two millennia.

(c) Theory and Practice

But if *Epidemics* 1 and 3 and *Prognosis* are not concerned with general physiology, other Hippocratic treatises certainly are. Their theories of the human constitution are in general recognizable adjuncts to later Presocratic pluralist physics (I.4): thus *On Regimen* makes fire and water

end p.58

the basic bodily constituents, constructing a complex account of the physiological and psychological characteristics caused by the various elemental blends.[7] On the other hand, in the defence of medicine's usefulness which opens the sophistically influenced public lecture *On Breaths*, the author explicitly contrasts the sort of surgical, interventional skill which may be obtained from practice with the more recondite theorizing necessary in other cases:

76to distinguish the most obscure and difficult diseases is a matter more for opinion than skill. . . . One of these things is this, namely what is the cause of diseases, and what the origin and source from which arise the affections of the body; for if someone knows the cause of a disease, he will be able to administer what is beneficial for the body.[8] (*On Breaths* 1. 18–26 Jones)

Air (or *pneuma*) is the source of life, in neo-Anaximenean fashion: any interference with its proper inspiration causes imbalances, diseases, and eventually death (1. 4–5: people can go days without food or water, but only a matter of minutes without air). Breathing bad air is responsible for epidemics (1. 6), while non-epidemic maladies are caused by poor regimen interfering with the proper functions of the body (1. 7). On the other hand,

77I consider that no constituent of the body contributes more to the intelligence than does blood. While the blood stays in its normal condition, intelligence remains normal; but when the blood alters the intelligence changes with it. There are many indications that things are thus: first, sleep, which is common to all animals, supports my contention. When it comes upon the body, the blood is chilled, as sleep is by nature chilling; when the blood is chilled it becomes more sluggish. This is clear: the body grows heavy and sinks (all heavy things sink by nature); the eyes close; the intelligence alters while certain opinions linger which are called dreams. (14. 4–19)

That passage is typical (the author goes on to describe and explain the psychotropic effects of drunkenness in similar terms) of the way in which vague empirical considerations are pressed into service as ‘indications’ for some general physical thesis. Epilepsy is due to air causing selective blockages in the veins, in some places stopping the blood altogether, in others speeding its movement, bringing about loss of consciousness, violent body movements, and foaming at the mouth (14. 31–54). Our writer concludes, with sublime confidence:

78 thus breaths are seen to be the most responsible causal factors in all diseases; all other things are auxiliary or contributory causes (*sunaitia kai metaitia*). This cause (*aition*) of diseases has been shown by me. I promised to enunciate the
end p.59

cause of diseases, and I have shown how wind (*pneuma*) is the active power not only in things as a whole, but in animals' bodies. I have tested the account against well-known maladies, in which the hypothesis has been shown to be correct. (15. 1–9)

The fundamental importance of air was also urged at this time (*c*.430 bc) by Diogenes of Apollonia, conventionally considered the last Presocratic cosmologist. Running against the prevailing current of the times, he was a monist; and for interesting theoretical reasons:

79 it seems to me, in a nutshell, that all existing things are alterations of the same thing, and are the same thing. This is perfectly obvious. For if the things that now exist in the world—earth and water and air and fire and the other things that clearly exist—if any of them were different from any other, being distinct in its own peculiar nature, and were not the same thing changed in many ways and altered, then they could not mix with one another in any way, nor could one benefit or harm another. (Diogenes, in Simplicius, *On the 'Physics'* 1. 4. 151. 31–152. 5 = 64 B 2 DK = 599 KRS)

Far from saving the appearances of generation, pluralisms actually make it impossible. Diogenes relies upon an abstract causal principle to the effect that

(CP4) for any x and any y, if x is to interact with y then x and y must be of the same type.

Versions of CP4 may be found elsewhere in the theorizing of the period (notably in Empedocles' theory of perception);[9] but I shall not pursue it any further here, beyond remarking that as it stands CP4 is too vague to generate any interesting or fruitful metaphysical consequences.

Diogenes was a physiologist as much as a philosopher; equally some of the Hippocratics were explicitly aware of Presocratic natural science. *On the Nature of Man* begins as follows:

80 for whoever is accustomed to hearing speakers discuss the nature of man in ways which go beyond its relevance to medicine this account will be of no use. For I do not hold that man is air, or fire, or water, or earth, or anything else that is not clearly a human constituent. I leave such accounts to those who like giving them. Those who do give them do not, however, seem to me to have
end p.60

proper understanding. For they all utilize the same idea, but they do not say the same things. They all draw the same conclusion from their idea, saying that there is one thing which is and which is unique and is everything; but they disagree about what to call it. One of them asserts that this one and all is air, another that it is fire, another water and another earth, while each adds to his account signs and indications which are of no import. (*On the Nature of Man* 1. 1–19 Jones)

The author points to the prevalence and irresolubility of the disagreements among the abstract theoretical physicists; their own arguments are self-refuting, serving only 'to establish the theory of Melissus' (1. 31–5). On the other hand, doctors tend to fasten on blood, bile, or phlegm as the basic human constituents, with the corollary that man is a kind of unity (2. 1–10). Our author disputes this, holding that if the human body were such a unity it could never feel pain (alluding to Melissus' deduction that his One cannot suffer pain since pain is a type of alteration: 30 B 7 DK). Furthermore, explicitly rejecting CP4, he holds that

81 generation cannot come to be from a unity; for how could something unified bring something about, unless it mixes with something? (3. 1–3)

Moreover,

82 all things are born in a like manner, and die in a like manner. For their nature is composed of all the things mentioned, and are resolved into the things from which they are composed; that is how they pass away. (3. 24–9)

The constituents in question are not the theoretical elements of the natural philosophers; rather, in line with the programme of **80** they are (allegedly at any rate) obvious bodily components: blood, phlegm, yellow bile, and black bile (4. 1–4). Health is the proper mixture of these constituents, while pain is felt when any one predominates at the expense of the others (4. 4–20).

His introductory protestations notwithstanding, our author's theory is itself abstract and removed from empirical observation (most notoriously in its introduction of the mythical black bile).[10] Although he attempts to

supply his views with empirical support (when people die, they secrete various different fluids; the single-element theorists must have been familiar with only one: 6. 1–41), the evidence, such as it is, evidently falls far short of establishing what he seeks to prove. Succeeding chapters attempt to integrate the four-humour theory into a general typology of the seasons, and of combinations of the basic qualities, hot, cold, dry, and wet: thus phlegm, being cold and wet, is associated with winter (7. 1–71). Here too the author adduces empirical evidence: 'an indication that phlegm is very cold is that if you touch phlegm, bile, and blood you will find phlegm the coldest' (7. 4–7); 'you may learn that winter fills the body with phlegm from the following considerations: in the winter sputum and mucus is fullest of phlegm, while in this season swellings are generally white and diseases phlegmatic' (7. 12–17). As the spring develops, 'phlegm remains strong, while blood increases; for the cold abates, and the rains come on, while the blood increases accordingly through the showers and the hot days: for these times of the year are by nature most like it, as it is warm and moist' (7. 17–23). In the summer bile starts to predominate, first yellow bile as the blood decreases in power, and then in the autumn black bile; while with the onset of winter, the bile 'becoming chilled decreases in quantity, while the phlegm increases again because of the abundance of rain and the length of the nights', i.e., presumably, the predomination of the wet and the cold (7. 24–8). These associations are grafted onto an allopathic theory of disease and therapy (7. 8–9), while different individual constitutions, as well as different climatic conditions, are responsible for the variation in disease patterns from individual to individual and from year to year.

The theory is comprehensive, and highly general—yet is intended to provide explanations in particular, individually distinct cases. Nor is it entirely unempirical: observations are adduced in its support (nosebleeds are more common in late spring and summer); and the supposed elements of the body are at least for the most part identifiable bodily components. But no effort is made to support the claim that they are *causally* basic in any real sense—the evidence at most supports their correlation with different seasons and their typical ailments—and in spite of the commonsensical tone of the opening paragraph (**80**), much of what follows is speculative and aprioristic.

And although several texts in the corpus do appeal to observation, and in some cases even primitive experimentation, such appeals are in general naïve and insufficient to establish their authors' contentions. The author of *Airs, Waters, Places* tries to distinguish different types of water according to their lightness and clarity, holding that water evaporated by the sun is the lightest and sweetest of all. In support of this, he draws attention to the

phenomenon of the saltiness of sea-water (*Airs* 8. 3–13 Jones), inferring (correctly) that the sun draws off only pure water. He elaborates:

83 it raises the finest and lightest part of the juice even from men. The following is the greatest indication (*tekmērion*) of this: whenever a man walks or sits in the sun wearing a cloak, the parts of his skin touched by the sun do not become damp, for it draws up each layer of sweat as it appears. But those parts covered by the cloak or anything else do become damp. For the sweat drawn and forced out by the sun is retained by the covering so that it cannot be dissipated by the sun. (8. 13–23)

However, he concludes, for no very obvious reason, that

84 for this reason rain-water goes bad more quickly than any other, and has a foul smell, because it is gathered together and mixed from many sources. (8. 25–8)

There follows a fanciful account of how the evaporated water is distilled into its components, rain being formed from the lightest:

85 such waters are naturally the best; but they must be boiled and purified if they are not to have a foul smell and produce sore throats, coughs, and hoarseness in those who drink them. (8. 48–51)

The line of thought is again obscure, if not contradictory: if the lightest and best water is driven off in evaporation, how can boiled water be healthy?[11]

Even so, the injunction to boil water is empirically justified: the writer of *Airs* clearly knows that stagnant water is unhealthy, and he tries to explain it in terms of his own pathology. Our own explanation for this fact would, of course, be utterly different in structure, involving micro-organisms undreamt of by the Greek theorists. Yet much of *Airs* is, in a sense, well- motivated. It seeks to explain epidemic disease in terms of some causal factors that might (in order to account for the selective incidence of such diseases) be restricted to certain types of place and time; and climate, air, and water are the obvious ones to fasten on. Moreover, given that many of the endemic ailments of the time would have been water-borne parasitic diseases, this account is not, as it happens, wildly off the mark. Equally prevalent were various forms of recurrent malaria; and they are seasonal, since they are carried by the anopheles mosquito, whose breeding-cycle is associated with standing water (although of course this fact was not known until the nineteenth century). And air quality is relevant to the incidence of the various pulmonary complaints whose descriptions we find in the clinical record, as are seasonal and climatic variations.

end p.63

The one thing Hippocratic pathology cannot well account for is the infection of one person by another, noted by Thucydides in his description of the plague: why should people (like the plague-doctors) fall sick simply as a result of proximity to an affected person? Some of these phenomena can be explained by proposing a generalized environmental cause which is liable to affect anyone in the vicinity; and some explanatory mileage can be got from the concept that both the original patients and the ones they affect are breathing the same contaminated air. However, it is striking that the pathology of infection and contagion is better accommodated within the old religious model, exemplified by tragedies such as *Oedipus*, where the gods infect the unclean with a polluting miasma which can be communicated to anyone with whom they come into contact.[12] Indeed, the very fact that the concept of infection was associated with the discarded supernatural paradigm may well have motivated partisans of an alternative view to ignore it; and nowhere in the Hippocratic corpus do we find any recognition of the phenomenon of contagion, no doubt for the excellent reason that it cannot be accounted for by humoral pathology.

(d) Hypothesis and Empiricism

Texts like *Airs, Waters, Places* and *On the Nature of Man* show theory outstripping any empirical research and observation upon which it may be based. Yet *Epidemics* is, in its sober and dispassionate observation, severely empirical and anti-theoretical. If it is true that the Greeks were in love with grand theory, ever ready to extrapolate great hypotheses from the flimsiest evidential base, still a countervailing distrust of theoretic castles in the air runs like a constant thread throughout the history of Greek natural science. The most striking exponent of this empiricism in the Hippocratic corpus is the author of the text known as *On Ancient Medicine*:

86 all those who try to speak or write about medicine, hypothesizing for themselves as an hypothesis for their argument the hot, the cold, the wet, and the dry, or anything else they want, narrowing down the principle of the explanation of diseases and death for men and hypothesizing one or two things in all cases, clearly blunder in much of what they say. (*On Ancient Medicine* 1. 1–8)

On Ancient Medicine reacts against the excessively theoretical tendencies its author discerns among his contemporaries; and it is a plea for a return
end p.64

to an earlier, empirically based medical practice. Medicine arose, he says, in response to natural human ailments; it has progressed by observing what remedies are effective in which cases; and if pursued in this spirit there is no reason why it cannot be brought to an even higher pitch of empirical effectiveness:

87for this reason I have not considered it [i.e. medicine] to be in need of some new-fangled hypothesis in the same way as the unclear and problematic things are, concerning which it is necessary (if one is to try to say anything) to rely on an hypothesis, as in the case of things in the sky or below the earth;[13] if someone were to state, perhaps even understand, how these things are, it would be clear neither to the speaker himself nor to his audience whether what he said was true or not. For there is nothing which can be brought to bear upon it which would give clear knowledge. (1. 20–7)

Medicine is a respectable empirical science; the hypotheses of the speculative natural scientists, as well as abstract theorizing about the constituents of the body, are inappropriate and useless for it.

Yet our author avails himself of assumptions which seem at first sight equally speculative:[14]

88(i) Indeed I know that it makes a difference . . . whether the bread be of bolted or unbolted flour, of kneaded or unkneaded wheat. . . . (ii) The powers (*dunamies*) of each of them are great and different from one another. . . . For man is affected and altered by each of these things in one way or another: (iii) and the whole of regimen is based on these things, for the healthy, the convalescent, and the sick. So there could not be anything more useful and necessary to know than these things, and how the first investigators, researching properly and with the reasonings appropriate to the nature of man, discovered things. . . . (iv) For they did not consider the dry, the wet, the hot or the cold, or any of these things either to harm or to be of any use to man, but rather they thought that the strength of each thing, being more powerful than the nature of man which failed to overcome it, did the harming. . . . (v) The strongest of the sweet is the sweetest, of the bitter the bitterest, of the sharp the sharpest, and there is an extreme of each of these internal things. For they saw that these things were internal to human beings, and were harmful to them; for there is in humans salt, bitter, sweet, sharp, astringent and insipid, and countless others possessing powers varying in number and strength. (14. 1–35)

Powers are later defined as 'intensity and strength of humours': (22. 3–5), which at first sight seems to be a strongly theoretical notion; moreover, people in olden times

89suffered many and terrible things on account of their strong and bestial regimen, eating raw and unmixed things which possessed great powers. (3. 21–4)

However, it is crucial to realize that these concepts are, for our author, empirically based. 'Power' here is simply the observable capacity of a particular type of food to affect, destructively, the body; and, he claims, it is an empirical fact that raw and unblended foods (**88**(i–ii)) are 'strong' in this sense:

90all the causes of pain can be reduced to the same thing, that it is the strongest foods which most greatly and most obviously harm human beings, whether they be healthy or sick. (6. 15–18)

The 'hypotheses' to which the writer objects, then, are not respectable, empirical causal suppositions: rather they are the grand and unfounded physical theses of the theoreticians. The method championed by *On Ancient Medicine* consists in establishing connections between recurrent phenomena. This makes medicine a fallible, Humean science:

91 you will not find any measure, either of number or weight, in respect of which one will ascertain precise truth other than the sensations of the body. And for this reason it is a task to find things out so precisely that only a few scattered mistakes are made. Indeed I strongly commend the doctor who makes only small mistakes. Perfection is infrequently seen. (9. 15–22)

The physician's task, then, is to ascertain what types of thing are harmful to the body, and in what quantities. Our author does indeed think that if some quality in a food is harmful in a certain concentration, then the greater the concentration the greater the harm. But this is not supposed to be an a priori conceptual truth: indeed he warns against over-confident assumptions of causal connection on insufficient evidence (9. 21).

The theoreticians hold that

92 if there is some hot, cold, wet, or dry which harms a man, then the proper medical practitioner should treat cold with hot, hot with cold, wet with dry and dry with wet; (13. 3–7)

but even if true, such a theory can be of no practical value. If someone of weak constitution consumes ‘strong’ foods (uncooked, unmilled grain and raw meat), the results will be harmful, perhaps even fatal. But, on the theoreticians' principles,

93 what remedy should be prescribed for someone in this state? Hot, cold, dry, or wet? Clearly one of these, since if the harmful agent is one of them, it must be countered with its opposite, as their theory has it. . . . [Here our author remarks that the sensible course would obviously be to moderate his diet, with baked bread, for instance, in place of raw grain, irrespective of such theories.] What

will we say? That hot things should be administered as he was suffering from the cold, or the opposite? . . . In preparing the bread, did he remove the hot, the cold, the dry, or the wet? That which has been exposed to fire and water, and compounded of many other things each of which has its own nature and power, has lost some properties, and has been combined and mixed with others. (13. 15–35)

Not only are theoreticians' claims that one of their favoured elements is responsible for what occurs empirically groundless, these hypothetical entities are superfluous, adding nothing of explanatory value to what can already be empirically determined regarding the effects of certain foods; moreover, the assignation of them (at least in their theoretical senses) seems entirely arbitrary (compare, however, the response of Theophrastus: V.3*a*, **204**). The genuinely useful property-categories (sweetness, bitterness, acidity, and the rest) are presumably to be taken as phenomenal (**91**): something will count as sweet just in case it tastes sweet. But, as deployed by the theoreticians, Hot and Cold do not correspond to ordinary phenomenal temperature:

94 I am at a loss as to how people who adopt this position and depart from this [i.e. the traditional] way to make the art hypothetical can treat human beings, since I don't think they have isolated anything that is in itself hot, or cold, or dry, or wet, and does not participate in any other form. Rather I think they have the same foods and drinks as everyone uses, but assign to them the hot, the cold, the dry, and the wet, since it would be pointless to enjoin a patient to be administered something hot. For immediately he will ask, ‘what?’ (15. 1–13)

Hot, cold, wet, and dry are used, then, either (*a*) in their ordinary sense (it *feels* hot), or (*b*) in some remote theoretical sense. If (*a*), then empirically speaking they are of little therapeutic import: temperature may sometimes, he thinks, correlate with disease; but he denies, on empirical grounds, that those connections are causal. In any case, the body tends spontaneously to correct for temperature variations (15. 16–19). But if (*b*), they are useless, since we gain nothing by assigning such 'theoretical' terms to the concatenations of evident properties which form the basis of scientific generalizations.

In spite of his consistent empiricism, our author allows himself to speculate on the internal constitution of the body: but only to the extent that such speculation is prompted by and responsive to evident empirical facts. He contrasts the properties of wine (which affects everyone similarly, at least if consumed in large quantities) with those of cheese (which is agreeable, even nutritious, to some, yet physically deleterious to others). Since the effects of wine-consumption are universal, they must be ascribed solely to a power in the wine (*aitios*: 20. 32). By contrast, however, the variable effects of cheese must be explained in terms of individual constitutional

differences (20. 25–47). But to say that is not to invoke mysterious internal, unobservable elements. We simply discover by experience which (empirically determined) types of people, subject to what regimen, suffer when they eat cheese, and which do not.

Furthermore, he allows causal speculation on the basis of observational evidence:

95 it will make a difference if he administers something hot and astringent or something hot and insipid (or equally the cold and astringent . . . or the cold and insipid). For I know that absolutely opposite effects result from each of them, not only in humans but also in leather or wood, and in many other things much less sensitive than humans. (15. 18–25)

The causal assumptions in play here are not entirely innocent, nor are they obviously compatible with the strictest empiricism. But at least his account is empirically motivated: the principal indicator of whether a particular substance will have a deleterious effect is its taste rather than its temperature; and he appeals to the evidently corrosive nature of astringent liquids (perhaps wineskins and mixing-bowls are the 'leather and wood' to which he refers). Human tissue is more delicate than wood and leather, but not radically different in type from them. It is thus reasonable to suppose that the human system is subject to such corrosions: indigestion appears to be one of them.

Finally, consider his enigmatic closing remarks:

96 if a humour that is sweet changes form not by admixture, but of its own accord, what quality will it first become: bitter, salt, astringent, or acid? Acid, I think. Thus the acid humour will be next least appropriate to administer where the sweet is the least appropriate. If someone is able to happen upon something by external investigation, he will always be able to choose the best of all; and the best is invariably that which is at the furthest remove from the unsuitable. (24. 5–13)

Spontaneous humoral alterations may occur, a fact which must be taken into account in making therapeutic inferences; moreover, we may deduce what will transpire inside the body by considering what happens outside it. He probably has in mind degenerative processes such as the souring of milk or wine turning to vinegar, which are (in the

appropriate sense) spontaneous (i.e. without obvious *prophasis*: **63**, **73**); and it is reasonable to suppose that similar processes go on within the body as well. There is nothing, in the strong sense to which he objects, hypothetical about such speculations; our author urges us to apply our everyday knowledge of the behaviour of particular types of juice to our therapy. It is often suggested[15] that our author is susceptible to the same sorts of objection as he
end p.68

levels against his theoretical opponents: that charge is, I think, without foundation.[16]

2. The Sophists

(a) The Nature of Sophistry

Contemporary with the developments in medicine which we have been discussing was the appearance in the Greek world of the Sophists. Sophists have had a bad name ever since Plato painted—perhaps caricatured—them in his dialogues as venal charlatans peddling a counterfeit virtue for money at the expense of truth (see e.g. *Gorgias* 448c–466a). Scholars have disputed the accuracy of Plato's portrait; that need not concern us. What does is the extent to which the intellectual current which they represented introduced new elements into the discussion of causation and explanation.

The sophists were self-proclaimed teachers; and pre-eminently they taught skill in argument, a very saleable commodity in the litigious climate of Athens in the late fifth century bc . This professional concern with forensic rhetoric lies behind the charges of ‘making the worse argument appear the better’, which Aristophanes levels at Socrates in the *Clouds* (882 ff.), and which Socrates in the *Apology* (19b) claims did so much to discredit him. And it is also at the core of Plato's charge against the Sophists: besotted with the charms of Persuasion, they are unfaithful to Truth.

That Platonic contrast is overdrawn; persuasion is not invariably deployed to disseminate falsehood. Scientists will seek to persuade people of the truth of their positions, and it is neither surprising nor intrinsically reprehensible if they employ rhetorical techniques to do so. If their intention is merely to conceal the frailties of their argument behind a rhetorical façade in order to dupe a gullible audience, the practice is disreputable; but there is no reason to think that it need always conform to that pattern (cf. *Gorgias* 457b–c; Aristotle, *Rhetoric* 1. 1. 1355^{a}21–^{b}15).[17]

But there is no doubt that the development of sophistic rhetoric served to draw attention to a series of crucial questions concerned with realism and objectivity. Sophistic thought tended towards relativism and subjectivism (which I shall not rigorously distinguish). Protagoras famously held that ‘man is the measure of all things: of things that are that they are, of things that are not that they are not’ (80 B 1 DK). While the interpretation
end p.69

of the ‘man-measure’ is controversial (crucially does he mean humankind in general, or individual humans?),[18] I shall assume (with Plato: *Theaetetus* 152aff.) that Protagoras means that everyone is judge and jury in their own case. If something seems to be *F* to *a*, then it is *F* (at least to *a*). Warranted judgement will thus have nothing to with objective truth, at least if that is construed as truth about a world independent of experience; and that has obvious consequences for the scope and structure of explanation.

Relativism is particularly plausible in the case of aesthetics: *de gustibus non est disputandum*. More importantly, if more controversially, the same may seem to hold in ethics as well. As sceptics were to point out with relish, moral values are culturally relative, and it is hard to see how disputes between values can be resolved without simply begging the question one way or the other. Perhaps ethical precepts simply express preferences, for which we may give detailed psychological aetiologies but no objective moral grounding.

All that is familiar enough—but it became so as a result of the sophistic temperament, and in particular because of the sophistic interest in the question what was owed to nature, *phusis*, and what to convention, *nomos*. If moral values are conventional, then nothing but conventional goods may be gained by adherence to them: the rational egoist will seek to circumvent them as and when it appears advantageous to do so. If I can get away with breaking the law, and do not, I am harming myself, since I am acting against the interests of my *phusis*, or nature. This is the burden of the surviving fragments of Antiphon the Sophist's *On Truth*; and it lies behind the views Plato ascribes to the Sophist Thrasymachus in *Republic* 1.[19]

The interrelation between these concepts is complex, and no brief account can do justice to them. But in so far as we are inclined to think of social mores as merely norms, chosen for convenience' sake, to that degree we are likely to think that sociology, psychology, and anthropology are the appropriate disciplines to which to turn for explanation of them; and equally we are likely to think (although this is not incumbent upon us) that human action is best described in terms of psychological drives and compulsions, and not of adherence to or deviation from the moral law. This in turn opens up large questions about the nature and extent of human responsibility, and the explanation of human action, which will be considered further in later chapters (IV.4*a–b*; VII.2*a–e*; X.4*c*; XI.5*c*).

(b) Responsibility and the Explanation of Action

I am responsible for my own actions. That seems trivially true: an event connected with me is an action of mine if and only if I am responsible for it. Yet the notion of responsibility is protean, involving both causal and moral components which are not always rigorously distinguished. We can be morally responsible for things we have not ourselves done:[20] we can inherit debts, and our children's misdemeanours can be laid at our door. Nor does causal responsibility entail moral responsibility; there are cases where even though I do something, I may not legitimately be blamed for it (if I am coerced, or if

I could not have been expected to foresee the consequences: see IV.4*b*). We may still suspect that causal responsibility has something to do with at least some indirect moral responsibilities (I am causally responsible for my children, which is what makes me morally accountable for their misdeeds); but even that seems too restrictive (I am not causally responsible, at least directly, for my adopted children). At all events the concept of responsibility is intimately linked with the explanation of human action; and the Greeks were the first to venture into the morass of difficulties and confusions associated with their relationship.

We possess, from the latter part of the fifth century bc , a collection by one Antiphon[21] known as the *Tetralogies*: three sets of forensic exercises pitting pairs of speeches by the prosecution against pairs for the defence. Some of them raise directly the issue of the link between causal and moral responsibility. Each deals with a different sort of homicide: the first premeditated murder, the second accidental manslaughter, and the third killing in self-defence.

In the second tetralogy, the subject is a case famous from Plutarch's *Life of Pericles* (36): one youth is throwing a javelin in a stadium, when another runs into its path and is killed. Pericles and Protagoras, according to Plutarch, spent an entire day debating whether the javelin-thrower, the games organizers, or even the javelin itself should be held responsible (*aitios*) for the death. In Antiphon's presentation, a further possibility is considered: the deceased himself may be responsible, through negligence, for his own death. The defence argues that the thrower was not wayward in his throw, and took all due precautions, while the victim did not (2. 2. 4–5, 4. 4–7); moreover, he was doing what he was supposed to be doing, and had been told to do (2. 2. 7). The defence case rests on counterfactuals: if the victim had not run into the path of the javelin he would not have been

end p.71

transfixed. By contrast, the prosecution relies on the defence's admission that their client was partially causally responsible for the outcome, and seeks to derive from that a measure of moral responsibility. Their case too turns on counterfactuals: if the youth had not thrown the javelin, the victim would not have been killed (2. 3. 10); hence he has a share in both the responsibility and the guilt for the slaying.

The prosecuting counsel paints the defence as maintaining that, in truth, the accused neither struck nor killed the victim (2. 3. 4–5, 7), which they say is absurd. What the defence actually said was: ‘the youth, while practising . . . made a hit, but did not kill anyone in real truth’ (2. 2. 3): the dead boy is himself at fault, and his culpability exonerates the accused of responsibility for what is in any case recognized on all sides to have been an accident (2. 2. 8).[22] Compare a claim made by a defendant in another of Antiphon's speeches:

> **97** moreover, involuntary mistakes are pardonable, while voluntary ones are not.[23] For an involuntary mistake, gentlemen, is the result of chance, while a voluntary one is caused by thought. (Antiphon, *On the Murder of Herodes* 92)

Antiphon's defence does not necessarily confound causal and moral responsibility when arguing that their client did not really kill his victim: they are making the altogether more sophisticated claim that (in some cases at least) an involuntary action is not to be thought of as action, properly so- called, at all.

The third tetralogy is also of jurisprudential interest. The defendant is accused of killing an older man by beating him; his defence is that the victim had attacked him. The prosecution argues that the victim died as a result of a beating so severe that it must have been intended to be fatal (*Tetralogies* 3. 1. 5–6). The defendant pleads that he reacted in self- defence to provocation (3. 2. 2). Indeed he claims not to have killed the victim, who survived for several days after the attack and only perished at the hands of an incompetent physician (3. 2. 3–4). His action was not premeditated, as murder must be (3. 2. 5). And even if his blows caused the death, they were themselves caused by the victim's original assault, which is thus ultimately responsible (3. 2. 6).

The prosecution replies that since the victim was elderly and infirm, it is false to claim that the defendant (who was in the prime of life) only retaliated in kind: his ‘weapons’ (his fists) were far more powerful than

those of his adversary (3. 3. 3). Moreover, the prosecution seeks to undermine the defence's claim to be merely the proximate, and not the original, cause of the death:

98 if our hands carry out what each of us intends, then he who struck without killing was the intentional author of the blow only; while the one who beat the other mortally was the author of the death. For the man died as a result of an intentional act on his part. (3. 3. 4)

The defendant cannot simply claim that his action was involuntary, since he intended to strike the victim (the defence counters that he intended to strike, but not to kill, thus drawing attention to the importance of the description under which actions are treated: 2. 4. 4–5; cf. IV.2).

The case introduces large issues of original responsibility, the right to self-defence, proportionality, and the *mens rea*. Its treatment of them is not always as lucid as it might be: but it demonstrates beyond question that such issues in the philosophy of law and of human action were already the subject of fierce debate in Athens in the late fifth century bc . Two further cases, possibly by the same Antiphon, are relevant to these questions. The *Prosecution of the Stepmother for Poisoning* and *The Death of the Chorus- Boy*[24] were both written for actual court-cases. In the former, a woman is accused of employing the concubine of a friend of her husband to poison his wine (claiming that it was a love-philtre), killing both her husband and his friend. The concubine was found guilty and executed; some years later a fresh charge is brought against the wife by her stepson. The prosecution claims that

99 in requital, the subordinate who carried out the deed got what she deserved, even though she was not really responsible (*aitia*) . . . the one who was responsible and conceived the plan will do so too, if you and the gods wish it. (Antiphon, *Prosecution for Poisoning* 20)

The adjective *aitios*, meaning ‘responsible’ or ‘blameworthy’, is the source of the nouns *aition* and *aitia*, meaning ‘cause’ or ‘reason’; and their initial connotation of intentional agency is gradually eroded as the term becomes established in a semi-technical vocabulary of causation. The evolution of legal theory and practice plays as important a role in this development as do the roughly contemporary advances in medicine (§1 above; esp. **67**, **75**, **78**).

One further legal development bears brief discussion. Both Presocratics

end p.73

and Hippocratics appealed to evidence (often of the flimsiest nature) to ground their claims; but it is in the law-courts that we discover the first conscious reflections on the nature and value of testimony as such. Of particular importance to this development is the refinement of the concept of an indication (*tekmērion*: cf. **83** above), a piece of evidence that leans in the direction of a particular explanation without entailing it.[25] Closely connected with this is the development of reasoning from probabilities: *to eikos*.[26] This concern with evidence and signs gradually becomes refined into a general theory of non-deductive, and hence defeasible, plausible inference—although it had to wait for Aristotle for that.[27]

(c) The Limits of Responsibility

It remains to consider Gorgias' *Defence of Helen*. Gorgias was a teacher of rhetoric, and *Helen* is a demonstration piece of flashy argument. But it has a serious core; and it raises fundamental issues regarding agency and the ascription of responsibility. *Helen* specifically concerns Helen of Troy's adultery: but the issues it raises are perfectly general. Helen has had a bad press at the hands of the Greek poets, getting the blame for causing the Trojan War. Gorgias rallies to her defence:

100 I want to give a reasoned account in my argument and so absolve her from responsibility (*aitia*). (Gorgias, *Helen* 2 = 82 B 11 DK)

He first offers a disjunctive analysis of the possible causes of her action:

101 she did what she did either (*a*) by the will of luck and the decision of the gods and the dictates of necessity; or (*b*) seized by force; or (*c*) persuaded by arguments; or (*d*) captured by love; (6)

and then proceeds to argue that in none of these cases is responsibility for what she did to be laid at her door. Gorgias thus implicitly (and reasonably)[28] supposes (*a*–*d*) to be jointly exhaustive: no action can fail to be comprehended under one or other of them (although some actions may have complex causes, and derive from more than one).

end p.74

The three conjuncts of (*a*) are treated as being similar in kind. One might rather expect the forces of necessity and of the gods to be species of (*b*); but in (*b*), Gorgias is thinking of human coercion and main force: and it may well have seemed natural to him to consider (*a*) as the cohesive class of all events for which no human can be assigned responsibility. The role of chance is unclear (indeed it is not clear what it means to say that chance is a cause at all: see further IV.2*b*). Even so, if an event really is fortuitous, then it appears that it cannot in any genuine sense be a human action at all, and *a fortiori* not one which may incur responsibility. We also surely accept that coercion and main force diminish responsibility, although we need not think with Gorgias that they are invariably completely exonerating (cf. IV.4*b*; VII.1*d*, 2*a*, *d*).

Case (*d*), the power of passion to derail the processes of common sense, was a Greek commonplace. The tragedians conceptualized sexual desire, *erōs*, as an alien force that fell upon an unwitting victim (e.g. Sophocles, *Antigone* 781ff.; Euripides, *Hippolytus* 525 ff.); and the more literally you take that account of the *psychopathia sexualis*, the more plausible it is to see passion not merely as being outside an agent's control, but as something which effectively destroys the possibility of genuine agency as completely as does physical compulsion. Gorgias' position relies on the plausibility of that account of desire. But even if we hold that our passions are a part of ourselves and not an external, alien force, we might still hold that we cannot be held responsible for them, since (we might argue) in so far as they are attributable to us their causes lie in our characters: but our characters are themselves the products of our upbringing, environment, and heredity, none of which are within our control (again Aristotle confronts the issue: IV.4*b*; cf. VII.1*f*, 2*a*–*c*, *e*).

That brings us to (*c*). Gorgias invokes the power of persuasion: words are charms that can bewitch the unprepared. But Gorgias' disjunction (*a*–*d*) clearly fails the test of exclusivity unless class (*c*) can include cases of self-persuasion; and indeed not just instances of self-deception, but all cases of rational deliberation. But then Gorgias must show not merely that words and reasons can influence an agent's behaviour, but that, in some suitably strong sense, they cannot belong to the agent at all: and it is hard to see how that can be done with any semblance of plausibility. Yet were Gorgias to adopt the slightly different line suggested at the end of the previous paragraph, he may perhaps still exonerate Helen (and everybody else). For even if my reasons for acting are *my* reasons, and hence not to be divorced from me as an agent, none the less for them genuinely to be reasons for action they must be systematically related with my desires. In general I have a reason to *F* just in case I like *F*-ing itself, or have a further reason to believe that *F*-ing is the best way available to me of achieving some other goal. Either way my desires have to figure in any account of my

rational action. But the structure of my desires, like that of my character, is the product of causal factors outside my control; hence I cannot be blamed for what I do as a result of them.

Gorgias' own considerations are less subtle, and will not make much headway against the ordinary human practice of ascribing responsibility. But they point the way forward; and within a century the issue of causal determinism and of its effects upon responsibility was to come to occupy the centre of the philosophical stage, whence it has never strayed far since.

(d) Sophistry and Science

Several of the texts of the **Hippocratic** corpus are sophistic in both style and content (*On Breaths* is a case in point). Moreover, some of them explicitly borrow the techniques of argument associated with the Sophists in order to argue that medicine is a real science, a *technē*,[29] and not, as its detractors allege, simply charlatanry. For quackery could not simply be ignored:

102 a good sign (*sēmeion*) that lay people are at their stupidest in discussing these [i.e. acute] diseases is that such cases give quacks their reputations as doctors. It is easy enough to learn the names of things prescribed to treat such patients, and if anyone talks of barley-water, or of a particular type of wine, or of hydromel,[30] the lay person thinks that all doctors, good and bad, mean the same thing. Far from it: in these matters they differ most particularly from one another. (*Regimen in Acute Diseases* 6. 4–15)

The doctors themselves are largely to blame for this state of affairs; and they are not interested in resolving their differences of opinion:

103 the science of medicine has fallen so low in the public view as not to seem to be a science of healing at all. As a result, if at least in acute disease-cases doctors differ so widely that a diet prescribed by one is regarded by another as harmful, they [i.e. the public] will say that the science has become virtually indistinguishable from divination. (8. 3–11)

Such sentiments are not uncommon: several treatises excoriate contemporary medical practice in the course of developing their own accounts of pathology and cure.

But such doctors must not allow medicine simply to be assimilated to divination; prognosis, so they think, should be a rational affair, based upon

end p.76

a scientific theory of the body's constitution. Consequently, they need to explain how it can be that a respectable science can be so uncertain in its results. This is the purpose of the sophistically influenced[31] treatise *On the Science of Medicine*, which begins by comparing, unfavourably, those who 'have turned the abuse of the sciences into a science of its own' with those who devote themselves to the difficult and slow task of accumulating knowledge. Its author then claims that 'there is no science which has no basis in fact', a contention he argues for in a notably confused manner:

104 it is absurd to suppose that something which exists does not exist—for what being could anyone assign to a non-existent thing to prove that it existed? For if it is possible to see non-existent things just as existent things, I do not see how anyone could think them non-existent, since they would both be seen by the eyes and known by the intellect to exist. Nothing of the sort is possible; what exists is always seen and known, while the non-existent is neither seen nor known. They are known when they are revealed by the sciences (*technai*), none of which is not established on the basis of a form. I hold that it is from the visible forms of things that they take their names. For it is absurd and impossible to think that forms sprout from names, since names are conventional, whereas forms are not conventional but are the offshoots of nature.[32] (*On the Science of Medicine* 2. 2–18 Jones)

Some sense may be sucked from that mire. The passage is Parmenidean in tone: reference to non-existent objects is impossible, hence scientists must be talking about *something*;

and even their arcane theorizings are the result of observations. Their theoretical terms are introduced in response to perceptible phenomena, and derive their meaning from the causal ancestry of their introduction. That has a strikingly modern sound to it: effectively it makes our author the Ur-father of the causal-historical theory of reference.[33] The final sentence asserts a robust realism: we cannot simply wish the world into existence on the back of our conceptual categories.

However that may be, the next paragraph helpfully attempts to clarify the point for the benefit of those 'who have not sufficiently grasped the argument' by moving from general to particular. Medicine has a subject- matter, and can be defined as

105 the complete removal of distress of the sufferers, the alleviation of the more violent diseases, and the refusal to undertake to cure cases in which the

end p.77

disease has already gained control, knowing that medicine cannot do these things. (3. 6–10)

Armed with this definition, our author opens his case:

106 the origin (*archē*) of my argument is granted by everybody: it is admitted that some who receive medical attention have been restored to health. But the fact that not everyone is cured is thought to be an argument against the science, while those who do recover from their illnesses, so the opponents of the science assert, owe their recovery to good luck (*tuchē*) rather than science. (4. 1–8)

He is perfectly prepared to allow that luck might play some sort of role (cf. *Places in Man* 46); but he insists that good care makes its own luck, and indeed later maintains that there is strictly speaking no such thing as a spontaneous recovery, since

107 every phenomenon has some cause, and if it has a cause, spontaneity (*to automaton*) is evidently no more than an empty name. Medical science is thus evidently real, both in regard to the causes of the various phenomena that occur, and in the provisions it adopts to deal with them. (*On the Science of Medicine* 6. 14–20)

The first sentence might be thought inconsistent with the theory of reference sketched in **104**: but it is not. The point is that we call certain events spontaneous when we are ignorant of their causes, and hence when we can neither predict nor intervene to influence them. But such events are still caused, which is why some people get better without medical care, a fact equally adduced by the opposition to prove the uselessness of medicine (5. 2–8). Such recoveries are to be explained, our author thinks, as being caused by the patient's regimen even if he is unaware of its beneficial nature (5. 12–25). Furthermore, the prescribed remedies may fail to be effective either because they are wrongly applied (science 'consisting in the discrimination between different procedures': 5. 30–5), or because patients refuse to adhere to them (7. 3–34).

None the less, a powerful objection remains to be tackled:

108 some blame the science of medicine because doctors are unwilling to tackle incurable cases. They hold that such diseases as the doctors do attempt to treat would have cleared up of their own accord, while those that need medical attention are neglected; while if the science existed it should cure all alike. (8. 1–6)

Other treatises emphasize the professional importance of not undertaking hopeless cases (cf. e.g. *On Prognosis* 1: §1*b* above); but our author explicitly gestures towards a

pathological theory of them. And he is surely right to insist that the fact that medicine cannot cure every case does not impugn its pretensions to scientific status, particularly if it can distinguish
end p.78

the hopeless cases from those where intervention is beneficial. On the other hand, he does nothing to establish that those who do accept medical attention are more likely to recover than those who do not:

109 in cases where we are given the tools by nature and by science to gain mastery, in these we can be practitioners; in the others we cannot. When someone is stricken with a disease more powerful than the tools of medicine, it cannot be expected that medicine will overcome it.[34] (8. 14–19)

He can provide an explanation of the failure of medicine to deliver in all cases; but crucially he does nothing to establish that it is the *right* explanation.

3. History

(a) Ethnography and Explanation

The *History* of Herodotus (fl. *c*.440 bc) is a vast, sprawling work, ranging from the accounts of the abductions of mythical heroines with which he opens, through the early history of Greece and Asia as handed down in the tradition, via a long excursus on the culture and practices of Egypt, to the final defeat of the Persian invaders at Plataea in 479 bc . But running through the myth, anecdote, and popular ethnography is a strong consistent thread. By tracing the relations between the Greeks and the Asians from the earliest times until the momentous clash between the great Persian Empire of Darius and Xerxes and the fragmented Greek world of Ionia, the islands, and the mainland, Herodotus seeks not only to document the long history of mutual antagonisms between the two peoples, but also to find an explanation for the astounding historical fact that the tiny Greek states, constantly at each others' throats and lacking material resources, managed to defeat the greatest empire in the world.

He opens with the following declaration:

110 this presentation of history is the work of Herodotus of Halicarnassus, and is done in order that what has happened to men may not be erased by time, and so that the great and remarkable deeds of both the Greeks and the Barbarians may not become dishonoured; but particularly to show for what reason (*aitia*) they fought with one another. (Herodotus 1. 1)

His method is, unsurprisingly, naïve. He concentrates on the actions and intrigues of individuals in order to explain the sweep of history; he seems charmingly credulous when it comes to reporting traditional stories; and he is not averse to discerning the large hand of divine fate lurking behind
end p.79

human affairs. When the Persians attempting to capture Potidaea are overwhelmed by a tidal wave, the Potidaeans, not unnaturally, put their deliverance down to the partiality of Poseidon—the Persians having profaned a temple of his—and Herodotus comments: ‘it seems to me that they spoke correctly when they said this was the cause (*aition*)’ (8. 129). Furthermore, at the battle of Plataea, Herodotus notes that, although the battle was fought right next to a precinct sacred to Demeter, no Persian dead were found on the sacred soil,

> **111** nor do any appear even to have set foot on it, while around the temple on unconsecrated ground, the greatest number were slain. I personally believe—if one may have beliefs about such mysteries—that the Goddess refused to let them in because they had burned her sanctuary at Eleusis. (9. 65).[35]

On the other hand, he does not believe everything he hears; and he is inclined to prefer natural over supernatural explanations for events, even when he retails the latter, if the former are available. It is not enough to invoke fate and oracles to explain the staggering reversal of the Persians at Greek hands (although he will on occasion do both):[36] what is required is some naturalistic account of how this could have come to be. Herodotus turns to comparative anthropology and ethnography, as well as to the study of political institutions. The Greeks are by nature sturdy and independent, made lean and tough by the exigencies of winning a living from their recalcitrant land. By contrast the peoples of the great Asian empires are soft and flabby, corrupted by easy living and luxury; and nor are their political institutions—tyranny and absolute rule—conducive to good morale. The Greeks are poor but free, and prepared to fight for that freedom.

Herodotus dramatizes the difference in a famous scene set prior to Xerxes' invasion. Xerxes sends for Demaratus, a renegade Greek, to seek his opinion on whether the Greeks will make a stand or submit without a fight. Demaratus replies:

> **112** poverty has always been the condition of my country; but she has the excellence that derives from wisdom and the strength of law. Relying on that excellence she wards off both poverty and slavery. . . . First of all there is no way in which they [i.e. the Spartans] will accept terms from you which would enslave Greece; they will carry on the struggle even if the whole of the rest of Greece capitulates. (7. 102)

Moreover, they will fight against any odds; for

end p.80

> **113** although they are free, they are not free in every respect: for they recognize one master, namely law, which they fear more even than your subjects fear you. (7. 105)

After the Persian defeat and retreat, Herodotus ends his history anachronistically but pointedly, by recounting a story about the emperor Cyrus. The Persians had not always been addicted to luxury and soft living: they too had derived from rugged, hardy, mountain stock. But when they achieved local hegemony they proposed leaving their harsh land and finding another more amenable:

114 Cyrus heard them, but did not think much of it; he told them to do so if they pleased, but ‘to prepare to be ruled rather than rulers; soft lands breed soft people; bountiful fruits and men excellent in war do not spring from the same soil’. Then the Persians saw that Cyrus' reasoning was the better, and they withdrew from his presence choosing rather to be rulers on a bare mountain than slaves in a well-tilled valley. (9. 122)

Large-scale explanations that invoke ‘national character’ may be simplistic and naïve. But they are perennially popular: and they may contain a kernel of truth.

(b) Ethnography and Medicine

A similar tendency to account for the varying characteristics of different peoples on the basis of their environment and institutions may be discerned in the second part of *Airs, Waters, Places* (12–24), an essay in explanatory anthropology only tangentially related to the climatic and geographical theorizing of the earlier chapters. The author begins by outlining his purpose:

115I intend to compare Asia and Europe, to show how they differ from each other in every way, and how the races of one differ from those of the other in physique. (*Airs, Waters, Places* 12. 2–5 Jones)

He contrasts Asian abundance with European poverty, explaining it in terms of comparative climatic factors, and concluding that ‘courage, endurance, industry, and spirit could not flourish by nature in such conditions . . . rather pleasure must rule’ (12. 40–4). He then sets out to document the physical differences among peoples:

116I will describe only the conditions of those which depart greatly from the norm, whether this be due to nature (*phusis*) or convention (*nomos*). (14. 2–3)

Most he attributes to factors of climate, regimen, or lifestyle; but he begins by essaying a Lamarckian explanation of the cranial characteristics of the ‘Longheads’:

end p.81

117no other race has heads like theirs. In the beginning convention was most responsible for the length of the head, but now nature reinforces convention. (14. 4–8)

This people, he explains, thought elongated heads a mark of nobility, and hence sought to promote this feature by means of binding in infancy. But after some time the characteristic became inbred:

118in the beginning convention was responsible, so that this nature came about as a result of force; but after a time, the process became natural, so that convention no longer compelled it. (14. 16–19)

The equability of the climate is chiefly responsible, he thinks, for the Asiatic lack of mettle (16. 3–14); but he does not discount the influence of political institutions:

119most of Asia is ruled by kings. Where people are not their own masters and independent, but are ruled despotically, their interests do not have to do with cultivating military virtues, but rather with how to give the impression of not being warlike (16. 16–21)

Slaves have no incentive to fight for their masters' interests.
Finally, the author devotes several chapters to the peculiarities of the Scythians. The dampness and coldness of their climate, as well as their fondness for horse-riding, are responsible, he thinks, for their deplorable physical condition (flabbiness, weakness, impotence, and lack of sexual desire: 16. 17–22). One case is of particular interest, that of the strangely effeminate Anaries, men who adopt women's roles as a result of incomplete sexual development (16. 22). They are seen by their fellows as divine prodigies (a view echoed by Herodotus: 1. 105); but our author, after asserting that all physical conditions have determinable physical causes, contends that their condition is caused by excessive indulgence in riding, which promotes impotence. The afflicted blame the gods: but the wealthy (who have more opportunity to ride) suffer disproportionately, while if the disease were divine it should affect everyone more or less equally (or possibly the poor more than the rich, since the rich can afford to propitiate the divinities: 22. 37–53). This recalls *On the Sacred Disease*:[37] and it provides a fitting epilogue to this account of fifth-century scientific demythologizing.

4. Conclusions

In a development echoing the Presocratic discovery of natural investigation, towards the end of the fifth century bc , there occurred a movement

away from traditional types of explanation of disease in terms of malign divine intervention and punishment, in favour of a naturalistic, physiological model of human pathology. Diverse as this movement was, its practitioners stressed careful empirical observation of the course and incidence of types of disease (Thucydides: §1*a*; *Epidemics*: §1*b*), and a rejection of supernatural explanations for their occurrence (*On the Sacred Disease*: §1*a*).
These theorists began to distinguish systematically between standing, internal conditions of the body which might predispose to illness, and the external, triggering causes of such illnesses (§1*b*), as well as stressing the importance of accurate prognosis, not least in order to substantiate their claims to scientific understanding. The Hippocratics differed regarding both the type and the extent of desirable theorizing, some (*On the Nature of Man*: §1*c*) holding detailed theoretical physiology to be a prerequisite of successful medicine; others (*On Ancient Medicine*: §1*d*) rejecting such ‘hypotheses’ in favour of a more austere empiricism; but all sought to isolate the causes of physical conditions, with a view, at least in the best cases, to being able to cure them, generally by way of applying the allopathic causal principle that ‘opposites cure opposites’.
Parallel with the medical developments come advances in legal theory and practice, which force a new focusing on the concepts of evidence and responsibility (§§2*a–c*); and the two strands come together in the attempts of sophistically influenced treatises like *On the Science of Medicine* to justify rational medicine's claims to scientific status in the face of its detractors (§2*d*). Finally, Herodotus' attempt to supply a large-scale historical

explanation for the Greeks' astounding triumph over the Persians in ethnographical terms (§3*a*) mirrors the concern of the doctors with the importance of climate and lifestyle in determining general traits of physique and character, as well as of susceptibility to disease (§3*b*).
end p.83

III Plato

R. J. Hankinson

1. Metaphysics and Causation

(a) The Inadequacy of Mechanism

120 This is no trivial question you raise, Cebes, since it entails a thorough investigation of the cause of generation and destruction. . . . When I was young, I was a devotee of that wisdom called natural science, since I thought it a great thing to know the causes of everything, why it is generated, why it is destroyed, why it exists. (Plato, *Phaedo* 95e–96b)

The speaker is Socrates, the non-trivial question, whether the soul is immortal. Socrates not unreasonably thinks that to answer it we must range widely in the metaphysics of generation and destruction, and he proceeds to give a brief sketch of the type of theory, cosmological, physiological, psychological, which had so infatuated him in his youth. However, so far from slaking his thirst for knowledge, the speculations of the natural philosophers left him convinced he knew even less than when he had started. For, in his innocence, he had thought that he knew that humans grew through the addition of bone and flesh from food and drink (96c–d), and further considered that

121 my opinion was satisfactory that when a large man stood by a small one he was taller by a head, and thus was a horse taller than a horse; even more clearly, I thought that ten was more than eight because two had been added. (96d–e)

However, he is now ‘far from believing I know the cause (*aitia*) of any of these things’; and part of the source of his perplexity, at least in the mathematical cases, is that one seems to be able to create two (i.e. a collection of two things) either by addition or by division:

122 nor can I any longer be persuaded that when one thing is divided, this division is the cause of its becoming two, since a moment ago the cause of its becoming two was the opposite. Then it was their coming together and one was added to the other, but now it is because one is taken and separated from the other. I no longer persuade myself that I know why a unit or anything else comes to be by the old method of inquiry, and I do not accept it, but rather have a confused method of my own. One day I heard someone reading, so he said, from Anaxagoras' book, saying that it is Mind that directs and is the cause of everything. I was delighted with this cause and it seemed to me somehow good
end p.84

that Mind should be the cause of everything. I thought that if this were so, the guiding Mind would direct and arrange everything in the best possible way. So if one wanted to know the cause of each thing, why it is generated or destroyed, one had to discover what was the best way for it to be, or be acted upon, or act. (97b–c)

In his excitement, Socrates looked to Anaxagoras to explain why it was better for the earth to be flat or round than any other shape, better for it to be centrally positioned, and so on.[1] But having purchased Anaxagoras' works, Socrates was bitterly disappointed:

123 I saw that he made no use of Mind, nor gave it any responsibility for the management of things, but mentioned as causes *aqr* and *aithēr* and many other bizarre things. That seemed to me like saying that Socrates' actions are all due to his mind, and then, in trying exhaustively to enumerate the causes of my actions, to say that the reason I am sitting here is that my body consists of bones and sinews, because the bones are hard and separated by joints, that the sinews are such as to contract and relax, that they surround the bones along with the flesh and the skin which hold them together, then as the bones are hanging in their sockets, the relaxation and contraction of the sinews enable me to bend my limbs, and that is the cause of my sitting here with my limbs bent. . . . but he would neglect to mention the true causes, namely that, when the Athenians thought it best to condemn me, it seemed best to me to sit here for this reason and better to remain and endure whatever penalty they order. For, by the dog, I think these sinews and bones would have been in Megara long since, impelled by my belief in what is best, if I had not thought it better and more honourable to endure whatever penalty the city ordered rather than escape and run away. (98b–99a)

Explanations, Socrates suggests, involve purposes. It is not enough merely to say *how* things operate; one must also say *why* they work the way they do: and the proper form of such explanations is in terms of how they best achieve their objectives by so doing. Any explanation that fails to meet these criteria is broken-backed. It is one thing to say that without the material arrangements of bones and sinews and so on the real causes would not be able to function as causes; quite another—and manifestly absurd—to say that those conditions are themselves causes (99a–b).

So far I have rendered *aition* and *aitia*[2] standardly as 'cause'; yet there are obvious difficulties with any such translation. First of all, the notion of a cause in modern English tends to connote activity: a cause is something which does something. That is reinforced by the predominant modern

tendency to treat causes as events, and to analyse causal sentences as involving relations between events.[3] Yet such connotations are misleading. As we have seen, to be *aitios* for something is to be responsible for it (II.2*b–c*), and responsibility is a broader concept than that of causation. Plato defines an *aition* quite generally as 'that because of which (*di'ho*) something comes to be' (*Cratylus* 413a); and that usage, in Greek no less than English, may cover both efficient and teleological explanations (although in some places Plato will reserve the 'because of which' formula for efficient or mechanistic causes: *Lysis* 219a–d).[4]

An *aitia*, then, can be anything which is referred to as being in some way explanatory of something. But Socrates implies that to call anything which is not a description of a goal or a purpose a cause is to commit a logical solecism:

124 this is a case of not being able to distinguish that it is one thing really to be a cause, quite another to be that without which the cause cannot ever be a cause; and this is exactly what most people evidently do, groping around in the dark, calling something a cause, giving it an inappropriate name. (*Phaedo* 99b)

This linguistic stricture does not cohere well with Plato's practice elsewhere. At *Hippias Major* 297a, for example, he describes the *aition* as 'productive' of its effect. At *Lysis* 218d, he distinguishes between what is done for the sake of something (i.e. for some purpose) and what is done because of something (i.e. as a result of some antecedent cause); indeed the two can be opposites—we seek medical help *because of* disease but *for the sake of* health. Plato generalizes in evaluative terms: that which is neither good nor bad (medicine)[5] is dear to the good (health) because of the bad and hateful (disease). Finally at *Philebus* 26e–27b, distinguishing between the various metaphysical fundamentals necessary for the creation of the universe, Plato refers to the fourth of them as 'the cause for all things that come to be of their coming to be'. He elaborates:

125 so the nature of the producer differs in nothing other than name from the cause, and the cause and the producer may rightly be said to be one. (ibid. 26e)

First appearances notwithstanding, these texts are not in conflict with the view that proper explanation is rigorously teleological in form. Plato is not concerned to deny that generation is a causal process—rather he is insisting that it be fundamentally explicable, and that explicability is something which can only be obtained by invoking intelligence and purpose. He does not reject ordinary causal accounts out of hand; rather he considers them deficient (cf. §4*b* below): mechanistic accounts can (perhaps) explain how things work, but they cannot give any account of why they do so.

(b) Forms and Properties

Having rejected pure, Mindless mechanism, Socrates confesses himself to be at an impasse, unable to proceed as he would like, and now unsure of what he had previously taken to be perfectly adequate explanations. He describes how, despairing of arriving at the truth directly, he adopted a *pis aller*:[6]

126 I set out in this way: hypothesizing in each case the theory which I judged to be the strongest, the things which seemed to cohere[7] with it I posit as being true, both about the cause and about everything else, and those which do not as not being true. (*Phaedo* 100a)

The hypothesis in question is stated in the next paragraph:

127 I will proceed by trying to show you the form of cause with which I have busied myself. . . . I hypothesize something to be Beautiful itself in respect of itself (*auto kath'hauto*), and something Good and something Large, and the same with all the rest. If you allow me these, and agree that they exist, I hope to be able to show the cause on the basis of them, and to discover the soul to be immortal. (100b)

end p.87

The 'Beautiful itself in respect of itself' is the Form of Beauty. Socrates proceeds to say that anything else which is beautiful is so because it stands in a certain relation, called 'participation', with the Form: and crucially this participation relation *explains why* any particular is beautiful to the extent to which it is so. Socrates no longer understands, he says, the other 'wise' explanations of the natural scientists, to the effect that something is beautiful because of its colour or shape (100c–d):

128 I simply, naïvely, perhaps simple-mindedly cling to this, that nothing else makes it beautiful other than the presence of, or community with, or however you want to put it, of Beauty itself. (100d)

This is described as 'a safe answer', namely that 'by the Beautiful, beautiful things are made beautiful' (100e).

This expression is ambiguous: the dative 'by the Beautiful' may signal agency (i.e. the Beautiful *causes* beautiful things to be beautiful); or it may rather point to a formal relation: beautiful things are beautiful because they possess a certain property, namely beauty. Plato is often accused of confusion on this score; setting out to discover causes to replace those of the natural scientists, sets of conditions for the instantiations of properties, he ends up with something altogether different, namely a formal account of what it is for those properties to hold. And even if that is a perfectly respectable project, Plato's version is open to the charge that it is entirely vacuous. If I am asked what makes a particular Renoir beautiful, Plato suggests, I should not talk of the colour, or the line, or the composition, or the brushwork—I should simply say that it is so because it has beauty in it: and that strays dangerously close to Molière explanation.[8]

It can be rescued from such vacuity by treating the Form of Beauty as a genuine object, one which is in some sense responsible for the inherence of the property in its instances; and that is exactly Plato's view of his Forms. Yet that is hardly liable to commend itself to anyone not already enamoured of Platonic metaphysics. Moreover, it leaves open certain proper causal questions. Even if we accept that Simmias is taller than Socrates because of the Tallness in him (102b–c), none the less we may properly ask how that has come to be the case—why should Simmias instantiate tallness to a greater degree than Socrates? Plato may correctly point out that 'because he is Simmias' is not a helpful answer to the question of why he is taller than Socrates; but his preferred alternative is hardly more illuminating. Furthermore, there seems to be no attempt to connect explanation

end p.88

with the good, which is what motivated Socrates' initial dissatisfaction with natural science in the first place.

But all is not yet lost. We need to examine more closely some of the considerations which prompted Socrates' general dissatisfaction with the sort of causes on offer. In the course of sketching in the reasons for his disillusion with physical science, Socrates advances a general logical requirement he feels any reasonable explanation must meet:

(P1) the same effect cannot be produced by opposite causes;[9]

addition and division cannot both be responsible for the existence of twoness. In the later discussion concerning the inadequacy of explaining someone's tallness or shortness as being ‘by a head’ (100e–101b), it becomes apparent that he wishes also to subscribe to

(P2) the same cause cannot produce opposite effects.[10]

The effects in question are, for the most part, the coming to hold of particular properties by particular objects (or sets of objects); thus P1 may be expressed as

(P1*a*) if x brings it about that y is F, then it is impossible that there be some z, such that z is the opposite of x, and z can bring it about that y is F.

Much depends upon the precise characterization of the notion of opposition to be found in the apodosis of P1*a*. If the examples of addition and division are characteristic, clearly opposition can amount to neither negation nor polar opposition. The notion seems best glossed by way of the concept of incompatibility: x and z are opposites just in case nothing can be both x and z (although something may be neither).

This formalization is inadequate, however, and it is preferable to switch from individual to predicate variables:

(P1*b*) if x brings it about that y is F in virtue of x's being G, then it is impossible that there be a z such that z is G^*, and z can bring it about that y is F in virtue of its being G^*,

where G and G^* are (at least) incompatible properties (i.e. it is not possible that there is an x such that x is both G and G^*). P2 can now be given a similar expression:

end p.89

(P2*b*) if x brings it about that y is F (in virtue of its being G), then it is impossible that there be a z such that x brings it about that z is F^*.

Indeed the principles may be simplified to refer only to properties:

(P1*c*) if G-ness is responsible for F-ness, then G^*-ness cannot be responsible for F-ness; and

(P2*c*) if G-ness is responsible for F-ness, then G-ness cannot be responsible for F^*-ness.

It is tempting to construe P1 and P2 as stating that causes must be respectively necessary and sufficient for their effects.[11] But they do not in fact assert any such condition—and it will be better not to assume at least without good reason that Plato meant them to.[12]

The move to properties allows a further refinement: P1*c* and P2*c* no longer imply that if something is to bring about a certain change in something else it must always do so. Thus one and the same *object* could be responsible for distinct (indeed contrary) changes in different things at different times, provided that we are sure to specify that it does so in virtue of its possession of distinct properties. This is surely the (admittedly rather confused) import of the discussion of relational properties at 102b–103a: it is not because he is Simmias that Simmias is taller than Socrates, but because of the tallness in him; equally his comparative shortness of stature in relation to Phaedo is not a direct consequence of his being who he is, but rather of the (degree of) shortness in him. He could change his relations of comparative stature with the others without altering his

identity.[13] But if so, then there is nothing essential to Simmias about those relations—and we require a different explanation for their holding than simply that he *is* Simmias.
We can now read this back into Socrates' dissatisfaction with the physical explanation of his being in jail. The physical arrangements of his bones and sinews no doubt have something to do with his sitting down—but not his sitting here rather than there, his being in Athens rather than
end p.90

Megara. Moreover (although Socrates does not explicitly say so), one might well think that his physical arrangements were the consequences of some further cause (in this case the result of his and the Athenians' decisions); and that in consequence the physicalists mistook the direction of explanation.[14]
Thus the alleged causes fail to meet P2*c*; and that implies that we need to look further for the genuine and determinative reasons why Socrates is where he is; which in turn suggests that a full-blooded explanation will be a sufficient condition of its explanandum, but need not entail it. Conditions may still intervene to prevent a cause from exercising its causal powers, and if cause and explanation are to be kept closely allied, that fact had better be respected. What matters is that the particular property or properties whose instantiation is appealed to in order to explain a certain state of affairs should not in any sense be explicative of its opposite. This does not mean that in any ordinary sense causes are indefeasible, since compatibly with this condition there may be cases in which an x has the appropriate property F to bring about G in y, and yet does not do so: we need, then, to distinguish (more rigorously than Plato manages) between causes considered as a class of objects which standardly (although not invariably) produce certain effects, and causes considered as the properties in virtue of which those objects exercise their causal powers when they do so.
Here we should introduce a distinction between explanation-types and -tokens. P1*c* and P2*c* are general, and hence best interpreted as type- conditions; however, in explanation we are frequently (indeed generally) interested in explaining particular events. We do not care so much about why people in general might choose to sit it out in jail rather than fleeing: we want to know why *Socrates* prefers to do so. Indeed, our interests being what they are, we tend to demand explanations for things just when they appear to be unusual. Socrates' fleeing requires no explanation; Socrates' preferring death does. Hence we frequently want to know which explanation-type is instantiated in this particular case.
P1*c* and P2*c* do not, considered as schemata for explanation-types, demand that G-ness is associated all the time or in all cases with F-type things. But its connection with F-ness is not so casual for it to be explanatorily null.
Socrates further invokes a third condition: it is ‘monstrous’, he says, ‘that something big should be made big by something small’ (101b); that generalizes to
(P3) if x is responsible for F-ness, x cannot be F *
end p.91

—a cause cannot be the opposite of its effect; which suggests (although again does not entail) that

(P4) if x is responsible for F-ness, x must itself be F,

which is equivalent to the first conjunct of PCS (I.3*c*). However P4, on the assumption that nothing can simultaneously bear contrary properties in respect of the same thing (which is explicitly endorsed by Plato at *Republic* 4. 436b; cf. *Timaeus* 62d), does entail P3, which suggests that Plato thought P4 basic and in a sense too obvious to mention, and wished to derive P3 from it.[15] If that is right, then Plato's purpose will have been to strengthen the view that it is not the object itself, but rather certain properties of it, that should figure in properly worked-out causal explanations.

And this coheres well with what follows. So far Plato has had Socrates offer only what he describes as 'simple-minded explanations', ones which, as we have seen, court acceptability at the risk of triviality. Plato now elaborates on the essentialist theme he has introduced: Simmias is tall not because he is Simmias, but because of the tallness in him; he could cease to be tall without thereby ceasing to be Simmias. However, this is not true of Tallness; nor, moreover, is it true of *Simmias*' tallness, the individual property-instance (or trope if you prefer):

129 it seems to me not only that Tallness itself never admits of being both tall and short at the same time, equally the tallness in us will never admit the short or allow itself to be overcome, but one of two things happens: either it flees and falls back at the approach of its opposite, the short, or it is destroyed at its approach. It will not admit of enduring and letting in smallness, and thus be the opposite of what it is. I may let in and endure smallness and still remain who I am, being this short man. But the other, being tall, can never dare to be small. (*Phaedo* 102d–e)

Alteration simply is the acquisition and loss of properties: the properties which are themselves the subjects of the alteration cannot remain through the change *in the changing object*. The logical form of change or alteration will then be something like this:

(A1) if x changes in respect of F, then there are times t_1, t_2 such that x is F at t_1 and x is not-F at t_2.

As it stands, A1 is inadequate, since it does not allow for alterations in the degree to which x is F, nor does it make the process continuous. Those requirements are easily secured, however:

end p.92

(A2) if x changes in respect of F, then there are times t_1, t_2 such that x is more F at t_1 and x is less F at t_2, and for any t_i, t_j, between t_1 and t_2, such that $t_i = t_j$, x is more F at t_i than at t_j.[16]

A1 allows (although it does not entail) that at some time x may fail completely to be F. It also captures in formal dress the intuition that underlay an earlier part of the argument (70d–71a), namely that generation is from opposites. It does not, however, say anything about F-ness itself, where it comes from or where it goes to. Thus it allows Socrates to deal with the objection that the new theory, elaborated in line with P3 and P4, contradicts the thesis of generation from opposites.

The logically protean nature of the preposition *ek* ('from') is to blame here. 'x comes to be from F' may mean that x comes to be as a result of F (or the F-ness in something); or it may simply mean that x, in becoming F, does so from a state of not being F (or at least being F to a lesser degree): cf. I.3*a*, *d*. We may label these the 'causal' and the 'material' senses respectively (and flag them with the appropriate subscript). Plato clearly sees the distinction:

130 you do not understand the difference between what we are saying now and what we said then. Then we said that an opposite thing came to be from its opposite thing; while now we hold that an opposite can never become opposite to itself, either in us or in its nature. (103b)

Plato's language is not as lucid as it might be—but by 'opposite thing' it is fairly clear that he means 'thing oppositely qualified', i.e. bearing opposite predicates. Plato sees that, at bottom, Parmenides' attempt to demonstrate the incoherence of change and generation depends upon a confusion between the causal and the material senses of 'from'. We can happily agree that

(P5) nothing comes to be (F) from $_C$ not being (F)

while allowing that

(P6) in every case of coming to be, something comes to be (F) from $_M$ not being (F)

is true, and significantly so (compare P5 and P6 with CP1*a* and CP1*b*: respectively, the principles that nothing comes to be without a cause and that nothing is generated *ex nihilo*; I.3*d*).

Having made this enormously important conceptual move to draw the teeth of the Parmenidean abolition of change, Plato tries to give more flesh

to his formal explanations, to render them less subject to the charge of triviality. At 103c–d, he makes the distinction between heat and cold on the one hand and fire and snow on the other; they are distinct, yet none the less stand in determinate relations to one another.[17] Fire is invariably hot; and snow invariably cold. Moreover, on the approach of fire, cold must either retreat or be destroyed, since fire is the bearer of the opposite property, heat, which cannot coexist with cold (103d–e). Thus, in addition to the Form of Heat there is something else which always bears the property, in this case fire. This allows Plato to sketch a non-trivial account of causal explanation which none the less remains faithful to the basic principles P2 and P3. Fire is never responsible for something's being cold; fire, being hot, is responsible for the heat in whatever it warms.[18] Now Plato takes it to be in some sense a truth of definition that fire is hot; and he has been taken to task for this (although with what justice is unclear).[19] But in any event, his position can be partially reformulated in terms of defeasible empirical hypotheses. Perhaps he does require that there be something which is under all circumstances both hot itself and heat-inducing, but this need not be fire as such: that it is is simply a highly confirmed general hypothesis. That Plato does not require his 'subtler causes' themselves to satisfy P1 is clear from his second set of examples. At 103e–105b he considers the relations that hold between Oddness itself and the odd numbers. Here the explanations involved are of a purely formal type (we are not concerned with the *generation* of the property 'odd' in things, if such a thing is intelligible): what is it that makes any particular three-membered set odd? Just that, Plato says, Three participates in the form of Oddness: nothing can be a triad and yet be even. Of course many

end p.94

other things (the number five, for instance) are also odd; hence, for the 'subtler causes' (or at least some of them), the objects that bear the formal properties (105c), no version of P1 holds (although it does, trivially, for the property itself).

This brings us to an issue that has been lurking in the background for some time. P1 looks as though it rules out of court disjunctive explanations, yet disjunctive explanations of effect-types are both commonplace and useful; and disjunctive explanations even of particular effects seem occasionally to be called for. Consider the following machine.[20] It is arranged so that every time a blue ball enters it, a black ball is emitted; however, if either a red or a green ball is inserted, it delivers a yellow ball. In the first case we have a one–one mapping of input and output conditions; but in the second, the mapping is many–one. Suppose we can see only the output: then, on the emergence of a yellow ball, we shall be entitled to say that its emergence was caused by the insertion of either a red or a green ball; and the causal law upon which this particular explanation rests will be disjunctive in form. Such a disjunctive explanation is prima facie incompatible with P1. And yet the example of three and oddness suggests that, at a certain level, Plato must tolerate disjunctive explanation: 'why is this set odd?' 'because it is three-membered; and any one- or three- or five- or . . .($2n = 1$)-membered set is odd'.

Nevertheless, there is a clear sense in which the last explanation is provisional only. For it is only if we can give a clear, rigorous, and non-circular account of the conditions for membership of the open- ended disjunction that figures in the explanation that we can be sure both that we have in fact isolated a coherent class and that it has the formal characteristics necessary to explain our instance. In the case of 'odd' this is easily secured: an odd number is any number not divisible by two. But that suggests that any disjunctive explanation must itself be susceptible of a higher-order explanation which will show why the disjuncts group the way they do; and that will be done by referring non- circularly to some property they all share in virtue of which they produce the same results.

Furthermore, this is just as plausible in the non-formal case of my ball- producing machine. We should search for some property the red and green balls have in common which accounts for their both being followed by a yellow ball (perhaps they are of a smaller gauge than the blue ball, and hence diverted by an internal sieve; perhaps there is an internal colour sensor designed to sift balls whose surfaces reflect wavelengths of greater

end p.95

than a given frequency—and so on).[21] In default of that we should have to declare that our observed regularity was itself inexplicable: perhaps we were in the grip of a massive coincidence. But then we would have *no* explanation, not even a disjunctive one, for why things turn out the way they do, and hence no warrant for any confidence that they will continue to do so. Explanation, then, is essential for grounding legitimate expectations; and we abandon the Platonic project at the expense of giving up our pretensions to explain at all.[22]

Of course there is no reason a priori why the world should be intelligible in this way. We cannot, as Plato sometimes seems to want to, simply infer its intelligibility from conceptual analysis, since it is always going to be in principle an open question whether our concepts really do latch onto the world in the way we think that they do. Perhaps the dream of intelligibility is just that: a dream. At best it must be an empirical hypothesis (or meta-hypothesis, since it is a regulatory condition on the establishment of hypotheses themselves); but it is one which we have reason at least to hope to be justified, since we can make sense of the world only to the extent to which it is true.[23] Moreover, Plato's natural teleology provides further support for it (§4*a* below).
end p.96

One final point needs to be made about Plato's theory here before we can turn to its application in the argument for the soul's immortality. The 'subtler causes' are such as invariably to possess the properties which they induce in other things: that is to say that
(P7) if *x* is a subtle cause for some property *F*, then *x* is invariably *F*,
which in turn entails, by a standard ancient equivalence,
(P8) if *x* is a subtle cause for some property *F*, then *x* is *F* by nature (i.e. essentially).
Neither P7 nor P8 seem compelling for our ordinary notions of cause, even as they apply to the transfer of properties: the hotplate on my stove warms the pan, but it is not invariably warm (even if, in a relaxed sense, one might say that it was warm by nature). Of course, the hotplate is hot when it is doing the heating; and so
(P7*a*) if *x* is a subtle cause for some property *F*, then *x* is invariably *F when it is bringing about F*-ness
will be true. But P7*a* is notably weaker than P7.
Commentators tend to accuse Plato of committing a fallacy here, a fallacy formally analogous to that generally diagnosed in the arguments of Alcmaeon and Parmenides (I.3*c*–*d*). But P7 applies only to subtle causes: perhaps Plato felt he had further reasons for adopting P7 at least in regard to this special class. The argument might run as follows: if *x* really is the proper cause of *F*, then *a* causes *F*-ness simply because it is *a*. But in that case *a* cannot fail to be *F*, since being an *F*-causer is part of *a*'s nature, and if *a* is an *F*-causer then *a* is *F* (by P4). Consequently, if we find that some *a* (e.g. snow), which is allegedly the subtle cause of some property (coldness), but does not invariably possess that property (we discover an instance of warm snow), then whatever else is true it is not in virtue of being snow that snow cools things; hence we should look for some other categorization of substance *s* such that it was invariably the case that *s* both brings about (under suitable circumstances) *F*-ness and is itself *F*. And that need not be a trivial stipulation; rather it may function as a methodological injunction to carry out empirical investigation. However, even if Plato is acquitted of fallacious reasoning on this score, it will be a further question whether he is entitled to think that fire, snow, and the like in fact meet the criteria for causal subtlety. I propose to approach this issue obliquely, by
end p.97

way of Plato's application of his causal principles to the argument for the immortality of the soul.

2. Causation and the Soul

(a) Alcmaeon's Argument Revisited

I begin not with the final argument of the *Phaedo*, but with its close cousin, Plato's reworking of Alcmaeon's argument from the nature of a self-mover (I.3*c*):[24]

131[I] (*a*) Every soul is immortal. For (*b*) what is always in motion is immortal, and (*c*) if what either moves something else or is moved by something else ceases motion, then it ceases living. [II] (*d*) Only that which moves itself does not cease moving, since (*e*) it does not abandon itself, and (*f*) it serves as a source and origin of motion (*archē kinēseōs*) for the other things which are moved. [III] (*g*) An origin is ungenerable. For (*h*) it is necessary that whatever comes to be comes to be from an origin, but that (*i*) an origin does not come to be from anything. For (*j*) if an origin came to be from something, it would no longer be an origin. [IV] (*k*) Since it is ungenerable, it is necessary that it be indestructible. For (*l*) if the origin is destroyed, it will be possible neither that (*m*) the origin comes to be from anything, nor that (*n*) anything else comes to be from the origin, if (*o*) everything comes to be from an origin. So (*p*) the self-mover is the origin of motion. (*q*) It is impossible for this to be destroyed or generated, otherwise (*r*) all the heaven and all becoming would fall into a standstill, and (*s*) it would never again have anything from which it could begin moving. [V] Thus (*t*) whatever is moved by itself is said to be immortal, and (*u*) no one will shrink from saying that this is the essence and formula of the soul. For (*v*) every body in so far as it is moved from the outside is unsouled, and (*w*) in so far as it is moved from within and of itself is ensouled, so (*x*) this is the nature of the soul. (*y*) And if things are so, and there is nothing other than the soul which moves itself, then (*z*) necessarily the soul will be ungenerable and indestructible. (Plato, *Phaedrus* 245c–246a)

The text is compressed, and interpretation is bedevilled by several cruces, most important of which concerns the translation of (*a*): the Greek is compatible with the 'collective' reading 'all soul is immortal'. This has encouraged some (notably Bett 1986) to treat 'soul' here as a mass-term: soul-stuff cannot be extinguished. However, as the argument seems clearly intended to establish *individual* immortality, and as no such general appeal to the permanence of soul-stuff will conceivably do the trick here, we
end p.98

should try to take the argument to be about individual souls distributively unless forced to do otherwise.[25]

(*a*) is the probandum of the argument; and (*b*) supports it. The ever- moving is immortal, at least in the cases in which movement constitutes life for something (this is the point of (*c*)). [II] serves to narrow down the range of eternal movers to those which are self-moved (*d*), by way of two further considerations: self-movers alone 'do not abandon themselves' (i.e. they cannot fail to instantiate the property in question: (*e*)); and they function as origins (*archai*) for the motion of other things (*f*). [III] supplies a conceptual argument against the generability of *archai*: if an *archē* came to be it could not, *eo ipso*, be an *archē*. [IV] infers, by way of empirical considerations, their indestructibility: if *archai* were destructible (but not generable), then sooner or later the universe would run down, since nothing else would be available to produce either them or anything else (*m*–*o*); and given an infinity of past time, all generation would already have ceased (*q*–*s*). Finally [V] establishes that souls are the appropriate candidates for self-movers of this type (*t*–*z*): hence souls are immortal (*a*).

The most serious problem for the distributive interpretation surrounds [IV], which is more naturally taken as referring to only one thing, namely the *archē* of whatever there is, or (as Plato is later to characterize it in the *Timaeus*) the World-Soul (*Timaeus* 34c–39d: §4*a* below; cf. *Philebus* 30a; *Laws* 9. 893c–896e). [IV] can be read distributively (if *archai* are destructible, ungenerable, and finite in number, then sooner or later all of them will be destroyed, and hence nothing will get moving); but even if it is not, Plato may simply be arguing that what applies to any *archē qua archē* should apply to all; hence if any *archē* is destructible all must be, and so even the World-Soul will be subject to destruction.

In any case, the empirically based conditions of [IV] are adjuncts to the main argument. The crucial section is [II], in particular clause (*e*),[26] which states a metaphysical principle. Here too there is a problem of interpretation. The word I translate as 'since', *hate*, may also mean 'in so far as'. This difference is not trivial: indeed it reproduces the distinction between the two ways of taking Alcmaeon's general principle (I.3*c*). Consider

(MP1) whatever moves itself, in so far as it does not abandon itself, does not cease moving;

and

(MP2) whatever moves itself, since it does not abandon itself, does not cease moving.

end p.99

MP1 merely asserts that a self-mover moves as long as it exists; MP2 makes the much stronger claim that any self-mover must move (and hence exist) for ever.[27] *hate* can, I think, have either sense; but it seems that (MP2) (where 'since it does not abandon itself' explains why the antecedent entails the consequent) is required by the argument, unless we are, after all, to convict Plato of fallaciously confusing the two. For why should Plato imagine MP2 to be compelling? I suggest that MP2 is not supposed to rest simply upon conceptual analysis: rather it is meant to be the conclusion of an investigation of the nature of causal connections and powers.

Archai can have, according to [III], no further origins on pain of not being *archai* at all; an *archē* of motion, then, will be causally autonomous with respect to motion. But then nothing can interfere with its motion—and so if it moves at all, it moves for ever. That argument does not commit a modal fallacy; but it does make some generous and unsupported assumptions about what it is to be an *archē*. We may distinguish different formulations of the *archē* principle. Consider

(PA1) $(x)(x \text{ is an } \textit{archē} \rightarrow \neg (\exists y)(y \text{ is causally prior to } x))$;

but to be an *archē* is to be an *archē for* something, which suggests something more like:

(PA2)$(x)(y)(F)(x$ is an *archē* of y's being $F \rightarrow \neg (\exists z)(z$ is F & x is F in virtue of z's being $F))$.[28]

Because PA2 restricts the idea of being an origin to that of being the origin *for some particular property* (motion, let us say), it has no tendency to conflict with the denial of PA1: one may hold that something is an *archē* in the sense of PA2 without being committed to denying that it is in some sense caused, or denying that it can, in some other sense, have had an *archē*.

Thus, one might say, any individual is an *archē* of their own actions: by PA2, in the case of some particular individual's action there is nothing other than the individual in virtue of which the action occurred (i.e. nothing else to which we may properly appeal as its genuine explanation). But that is not to imply that individuals themselves are uncaused; their parents are causally responsible for them. To put it another way, to be an origin in this sense is to be an initiator; to be an initiator *O* for some sequence *S* is just for there to be no individual *I* prior to *O* in *S*. But that is

end p.100

quite compatible with there being any number of individuals prior to *O* that are not in *S*. Once we see that there is no such thing as being an initiator *tout court*, that to be an initiator is to be an initiator for some sequence, there is no reason to imagine that initiators cannot themselves be created.

(b) The Final Argument of the *Phaedo*

The theory of explanation sketched by Socrates in *Phaedo* 95a–105c is a preliminary to another argument for the soul's immortality. 'What is it which is such that, when it is present in a body, it will be alive?' Socrates asks (105c), and he receives the predictable answer 'the soul'.[29] The argument is relatively straightforward:

(1) the soul is such that it brings life to whatever it occupies;
(2) death is the opposite of life;
(3) the soul will not admit the opposite of that which it invariably brings;
hence
(4) the soul will not admit death;
but
(5) whatever will not admit death is immortal;
so
(6) the soul is immortal (105d–e).

(1) assumes that the soul is something substantial which injects life into any body it occupies, a perfectly natural Greek view, but one which was not to go unchallenged.[30] (2) is taken to be an obvious truth of definition. The crucial premiss (3) rests on what has been established earlier in the argument, namely P7 and P8: soul is a 'subtle cause' of life. None the less, that argument offered two distinct possibilities as to what might happen under such circumstances—either the original property must be destroyed or it must retreat. It seems possible, then, to construe (3) and (4) in such a way that they only license the conclusion that a soul will not admit death *and still be a soul* (in line with P7*a*

above). But on that reading, (4) does not entail (5); and so the argument fails. Plato is aware of this (which makes it
end p.101

implausible that he simply commits the fallacy of confusing P7 with P7*a*); and he tries to draw the sting of the objection, by contrasting the properties of snow, fire, and the odd, with those of the soul. Fire is not conceptually indestructible; yet the soul is, since that which always brings life (106b–d). Thus he argues that (6) must entail that

(7) the soul is indestructible,

which supplies the conclusion to the argument. So (6) does not on its own entail the soul's immortality (in the sense of its eternal existence): that requires further argument. The problem is to see what that argument is.[31]

My suggestion is to read the argument in line with my treatment of the *Phaedrus* demonstration. Soul is a strong reason for life; consequently, in order to explain why something is alive one need go no further than point to the fact that it is ensouled. But then there can be no further explanation of anything's being alive; hence nothing other than the soul is causally or explanatorily relevant to life; but if there is nothing independent of the soul which has any such causal relevance, nothing can have any tendency to affect the soul's life-giving nature; and hence (by an application of P4) the soul cannot cease to be alive. That argument is confused; and the confusions are generated by an insufficient attention to the distinction between property and property-bearer, as well as by an uncritical adherence to the principle that causes must resemble their effects. If explanation is to be both entirely general and exceptionless, and is to proceed by way of identifying relations between properties, then those relations will, in the appropriate sense, be eternal. But that does not show that any *bearer* of those properties, even if in some sense it bears them essentially, will also be eternal. But to produce a proper analysis of this would take us deeply into the metaphysics of substance, something we must defer until the next chapter.

3. Hypothesis and Explanation

(a) *Meno* and *Phaedo*

Earlier, we saw Socrates carry an investigation forward by means of a hypothesis (in this case that there was such a thing as Beauty-in-itself, and so on); but a discussion of quite what the method entailed was deferred.
end p.102

Hypothesis is, however, closely linked by Plato to explanation: and as such we cannot defer the issue, thorny as it is, indefinitely.[32]

In the *Meno*, Socrates at first holds that Meno's question ‘Is virtue teachable?’ can only be answered on the basis of a knowledge of what virtue really is: by way, that is, of a real definition of virtue. However, baulked in their search for such a definition, he retreats and proposes to approach the problem indirectly, ‘by means of a hypothesis’ (*Meno* 86e). He then gives a famously obscure geometrical illustration of what he means (86e–87c). But however precisely that is supposed to be construed, it seems reasonably clear that a hypothesis in this sense has something like the following form. Suppose you want to know whether *p*; *p* is not susceptible of direct investigation, but there is a further proposition *q* which is so susceptible, and is such that if *q* is true, *p* is true (and, Plato suggests, vice versa): hence by establishing the truth-value of *q* we may establish the truth-value of *p*.

It is not clear which component here is referred to as the hypothesis; nor are the precise logical relations between hypothesis and probandum strictly delimited. Plato needs them, in at least one case, to be biconditional, yet all he apparently establishes is a one-way conditional: in the *Meno*, he proceeds by hypothesizing that if virtue is teachable, then it must be (a form of) knowledge (since only what is known can be taught: *Meno* 87b–c); and then seems to infer that since virtue *is* (a form of) knowledge, it must be teachable (87c). It is not clear whether this is simply a logical mistake, or whether Plato thinks he has an argument for the equivalence of knowledge and teachability (certainly the claim that all knowledge is teachable is less intuitively plausible than its converse), or whether he simply treats it as yet another hypothesis whose consequences deserve investigation (at all events, the theory of recollection commits him to the view that not all knowledge is to be won by teaching in any ordinary sense).

Whatever the truth of that, the hypothetical procedure has something in common with the geometrical method known to the Greeks as analysis.[33] Analysis in this sense is a form of reduction. You begin from some problem (an alleged theorem), and try and show what needs to be the case for it to be true; then what needs to be the case for *that* to be true; and so on until you arrive at an axiom, some (allegedly) self-evident truth. Then proceeding in the other direction you prove the putative theorem really to be a theorem by deriving it from the axioms (this is what Plato apparently has in mind when he introduces the geometrical example in *Meno*). But analysis is not simply a matter of reducing a particular problem to a higher- order theorem, and showing how it is simply a special case of a more

general truth: rather it is supposed help to *ground* those higher-order truths. That is why it is a two-way procedure, and why the entailments involved must go in both directions. It is one thing to show that a right- angled isosceles triangle exhibits Pythagoras' theorem because it is a special case of a right-angled triangle, and Pythagoras' theorem holds for all such triangles; it is quite another to argue that the theorem should be adopted universally because it holds in this case.

For Plato's purposes, two things matter. The first is that hypothetical arguments involve hierarchies of related propositions; and secondly that at some point in the ascent up the hierarchy, we must abut on a proposition the truth of which is not itself hypothetical. This aspect of the hypothetical method is evident in *Meno*: having established (he thinks) that if virtue is knowledge, then it is teachable, Plato asks the obvious question, namely is it knowledge (87c)? He proposes to answer that by considering the consequences of a further 'firm' hypothesis, namely that virtue is good; and he deduces from the premiss that (the exercise of) knowledge is what makes good things good that virtue must involve knowledge. We need not bother with the evident logical inadequacies of the argument as it is presented. What matters is the appeal to a higher 'hypothesis': and the related facts that it is both supposed to be a truth of definition, and that it appeals to the good. After all, that is where Socrates set out from in his search for genuine explanations. Furthermore, when he introduces the concept of hypothetical investigation, he says:

132 when you have to give an account of the hypothesis itself, you will do so as follows, by hypothesizing another hypothesis, which seems to be the best of those above, until you arrive at one which is sufficient. (Plato, *Phaedo* 101d)

Plato seems to envisage a process of ascent via new, more basic, and usually (although not necessarily) more general hypotheses until we discover one which is 'sufficient' presumably both in that it generates no inconsistencies and in that it represents some ultimate principle of explanation. For Plato, this will involve the notion of the Good. The ultimate hypothesis in *Meno* was that virtue was good; the methodological strictures of the *Phaedo* direct us towards explanations that refer to the Good. Good is of course an evaluative concept; and as such it figures in any general purposive explanation that invokes choice and deliberation: which leads us to our next considerations.

(b) Hypothesis in the *Republic*

Plato clearly distinguishes things which are pursued for the sake of their derivative benefits from those which are pursued for their own

sakes (*Republic* 2. 357b–d). Moreover, he stresses that the process of explaining goals in terms of further more basic goals must come to an end somewhere:

133 medicine is dear to us for the sake of health. . . . and health is also dear . . . and if dear then dear for the sake of something . . . and this object must also be dear. . . . then proceeding in this manner we shall arrive at some first principle (*archē*) of dearness which is not susceptible of being referred to any other, for the sake of which everything else is dear, and having arrived there we shall come to a halt. (Plato, *Lysis* 219c–d)

The idea of ascending to an *archē* which somehow requires no justification is expressed in a more abstract manner in a famous passage of the *Republic*, the image of the divided line. In it, Plato subdivides the activities of the intellect into two classes

134 in such a way that in one part the soul, making use of what was represented to it as images, is forced to inquire from hypotheses, moving not towards the *archē* but towards the conclusion, while in the other, which leads to an unhypothesized *archē*, the soul moves from a hypothesis, but without the images of the first part, but with the Forms it achieves its object by means of them. (Plato, *Republic* 6. 510b)

Plato tries to clarify this by way of an illustration from geometry. Geometers, he says, proceed on the basis of there being such things as odd and even, plane figures, different types of angle, and so on. But they simply hypothesize them, giving no account of their existence, or justification for their use of them, and then proceed to draw their conclusions (6. 510c–d). Moreover, they make use of visual representations of triangles and so on, although in a perfectly clear sense the theorems they prove are not *about* those representations (6. 510d–e: cf. **134**). This procedure does not allow the geometer to 'travel upwards from the hypotheses' to something firmer (6. 511a).

This ascent is to be accomplished by a second function of the intellectual faculty which Plato calls dialectic:

135 it does not treat the hypotheses as though they were first principles, but properly as hypotheses, as stepping-stones and starting-points, in order to reach the unhypothesized, the *archē* of everything, and having grasped it, and keeping a grip on what follows from it, it descends to a conclusion making use not of anything perceptible in any way, but of the Forms through themselves to themselves, and concludes with Forms. (6. 510b–c)

To treat hypotheses 'properly as hypotheses' is presumably to respect their provisional status. Plato rightly points out that a conclusion is only as good as the assumptions it rests upon: if they are unfounded, so too is the conclusion.

end p.105

Plato thinks that it is only by being able to ascend[34] to the unhypothe- sized first principle that we can proceed to rebuild the structure of our knowledge on firm foundations. His purpose then is akin to that of Descartes, even if both his method and his chosen starting-points are rather different. But crucially the structure thus exhibited will *explain why* things are the way they are *in virtue of* their fulfilling a role in the overall hierarchy. The relations between explanation and the structure of knowledge are thus intimate: it is only if the separate sciences can be shown to depend upon this unhypothesized first principle that their subjects can be truly known (6. 511c–d).

Plato does not work this out in detail; instead he supplies, in the celebrated image of the cave, a model of how the ascent takes place. Beginning from the objects in the sensible world, it somehow transcends the in such a way that their manifold and incurable indeterminacies and defects do not compromise the truth arrived at at the end of the journey. It is clear that Plato is groping for what a later epoch was to call transcendental argument. It is equally clear that he has no clear grasp of what any such argument must involve in order for it to be successful. But the vision of knowledge as consisting in the establishment of a secure hierarchy of explanation, all of which ultimately depends upon the principle of the Good, was to prove irresistibly seductive to later thinkers (XII.2*a*–*b*, 3*b*).

(c) Explanation and Understanding

Elsewhere, in passages that have nothing to do with the method of hypothesis, Plato seeks to emphasize the systematic connections between explanation and understanding. The idea that knowledge (or understanding: *epistēmē*) properly so called is of what is permanent and unchanging (and is to be contrasted as such with opinion, which is of the mutable, sensible world) had a great grip on the Greek mind. Plato enunciates the view classically in *Republic* 5. 476e–480a; and Aristotle too (although in a somewhat different sense) endorses it: ‘everyone supposes that what we have *epistēmē* of cannot be otherwise than it is’: *Nicomachean Ethics* 6. 3. 1139^{b}19–21 (cf. 6. 6. 1140^{b}31 ff.; *Posterior Analytics* 1. 2. 71^{b}10, 1. 4. 73^{a}22 ff., 1. 6. 74^{b}5 ff.; V.1*a*).

This sort of knowledge is closely associated with explanation. At *Meno* 97e–98a, Plato distinguishes between knowledge and true belief on the grounds that the former is firmly anchored by an account of the cause or

end p.106

reason why (*aitias logismos*) things are the way they are.[35] At *Theaetetus* 200e–201d, Plato again argues that true opinion and knowledge come apart, otherwise jurymen who are convinced by a clever but spurious argument for what is in fact the truth will be allowed to know that truth. Equally for Aristotle to have *epistēmē* or understanding is to have *explanatory* knowledge, to know *why* it is that a certain eternal truth obtains (*Ethics* 6. 6. 1140^{b}35; *Posterior Analytics* 1. 2. 71^{b}10 ff.: V.1*a*). Moreover, Plato regularly distinguishes a *technē* (art, skill, craft, science, technical ability, expertise)[36] from lesser types of competence such as *empeiriai* (empirically derived abilities) and mere knacks or *tribai*, on the grounds that *technai* have explanatory backing (*Phaedrus* 260e, 270b; *Gorgias* 462b–465a, 501a; for the opposing view, see Polus' claims in *Gorgias* 448c; IX.2).

Tying a belief down to a causal account serves a mnemonic function: Plato believes that genuine knowledge is stable and can never be lost, in contrast with labile belief. This is why the slave-boy's freshly elicited true belief in regard to the problem of doubling the square cannot as yet count as knowledge (*Meno* 85c–d); he must be able to repeat all of the steps of the proof himself, and come to see that it holds not just for arbitrary particularly instances, but for all relevantly similar cases. To have understanding of some proposition is to see where it fits into the general structure of things; and that is to be aware of its explanation. To have genuine understanding of a domain, then, is to grasp the general, explanatory truths that govern its articulation (here again Plato's views anticipate *in nuce* the more formally adequate accounts offered by Aristotle: IV.1*a–b*; V.1*a*); the crude generalizations supplied by experience alone are not enough to constitute a proper, respectable *technē*.

Plato's main test-case is medicine. Does it consist simply in a concatenation of experientially derived and grounded beliefs, or does it rather require a profound and detailed understanding, not merely of human physiology, but of the causal structure of the world as a whole? That was already a subject of Hippocratic debate (II.1*c–d*); and Plato himself invokes Hippocrates' name in this context:

> **136** consider, then, what both Hippocrates and the true account have to say about nature. Surely we ought to think about the nature of anything in the following way: first we must determine whether what we wish to be expert in (and to be able to make others expert in) is simple or complex. Then if it is simple, we must investigate what sort of natural capacity it has for acting and being acted upon, and upon what and by what. But if it has many forms, we

end p.107

must enumerate them and observe in each case what we observed for the one, namely what it naturally does to what, and how and by what it is acted upon. (Plato, *Phaedrus* 270c–d)

The holism of the procedure is further emphasized by a text in *Charmides* (156b–c), where Plato notes that when a patient visits 'eminent doctors' with an eye complaint, they will not treat the eye alone, but the head, and indeed the whole body, since 'the part can never be well unless the whole is' (156e). Thus medicine, here treated as a paradigm of a respectable science (cf. II.2*d*), one which is able to deliver genuine explanatory understanding of its subject-matter. In *Gorgias*, Plato develops a complicated account of the relations between real and counterfeit sciences; as so frequently, this takes the form of contrasting appearance with reality, false with true coin. Cookery is a sort of flattery, standing in the same relation to medicine as fashion does to gymnastics (*Gorgias* 465b); the one merely appears to tend to the body in a certain way, while the other genuinely does so (464b–d):

137 so I call it blandishment, and hold it to be shameful, Polus—for I'm saying this against you—because it aims at pleasure rather than what is best. I say that it is no *technē*, but rather an *empeiria*, since it can offer no account based on the nature of things as to why it prescribes what it prescribes, and consequently cannot speak of any of their causes. (464e–465a)

The same concerns may be found in the *Laws*, where Plato explicitly contrasts the two sorts of medicine. The first is suitable only for the slavish (since it is learned from masters, and not from 'nature itself'), and relies merely on experience 'as if it were perfectly accurate' (*Laws* 4. 720b), refusing to offer a causal account of the disease in question, in contrast with the other which does so 'scientifically, and from its origins, investigating its nature' (4. 720c), speaking 'almost as a philosopher would, tracing the illness to its origins, and discoursing on the whole nature of the body' (9. 857d). Understanding, then, is understanding of the whole; and it is to the latter project in its most general form that Plato devotes his *Timaeus*.

4. *Timaeus* : The Model Applied

(a) The Basis of the Argument

The *Timaeus* is in many ways a curious dialogue. In fact it is barely a dialogue at all: rather it is a monologue on physics and cosmology put into the mouth of one Timaeus, supposedly an adept of astronomy (*Timaeus* 27a). Timaeus proposes to give an account of both the structure and the generation of the cosmos, but the account is not offered as scientifically certain: it is merely a ‘likely story’, an *eikōs muthos* (29d: this emphasis on likelihood rather than certainty is repeated throughout the dialogue). The reasons given for this are obscure:

138 an account of what is stable, fixed, and clearly intelligible will itself be stable and unalterable: in so far as words can be irrefutable and unshakeable, it will not fall short of this. But accounts of what is created in its likeness, being of a likeness, will be merely probable in the way the likeness is; for as becoming stands to being, so belief stands to truth. (28b–c)

First appearances notwithstanding, Plato does not intend the creation story merely to be a myth, a way of accounting for the physical world in a sense, but not one which is meant to carry conviction. The account will show how a divine craftsman, the Demiurge or Artisan, created the visible world as a representation or likeness of the perfect paradigm forms in his mind. The point of **138** is that, in so far as the sensible copies of those paradigms fall short of their originals, we are unable to infer with certainty their precise structure. We may deduce that the Artisan acted for the best—but we cannot determine fully how that intention was carried through.[37]

The Artisan's goodness is asserted immediately afterwards:

139 let us say for what reason the constructor constructed becoming and the universe: he was good, and what is good has no trace of envy in it; being thus without envy, he wanted all things to be as much like himself as possible. One may accept this as the most powerful and correct principle (*archē*) for generation and the cosmos available to thinking people. (29d–30a)

So since order is preferable to disorder, the Artisan created ‘as far as possible’ order out of the chaotic ‘disorderly motions’ that preceded the imposition of structure on the world (30a: cf. 52d–53c).

‘As far as possible’ is crucial here. Plato's Artisan is not, like the God of Genesis, a creator *ex nihilo*. He is an organizer, working with pre-existing materials which impose material constraints upon what he can produce. Limits are set on the range of creative possibilities open to him by the intransigence of matter. This limited teleology has clear advantages over its Christian rival, as it renders the Problem of Evil (in its most general form: if God is good, why aren't things better than they are?) less of a problem. The more circumscribed the creator is by material exigencies, the less the imperfections of the world reflect badly upon him. Thus the standard of goodness to which the world has to aspire is less stringent than it is

for a creation *ex nihilo*, and the teleology is consequently less empirically implausible. The implicit rejection of creation *ex nihilo* represents an acceptance of the Parmenidean principle CP1*b* (I.3*d*); that he equally admits CP1*a* (the ban on causeless generation) is evident from Timaeus' preliminary speech:

140 everything that is generated or changed must be so because of some cause, since it is utterly impossible for there to be generation without cause. So whenever the Artisan looks to what is eternally unchanging, making use of a paradigm of this sort, and reproduces its form and capacity, then he must necessarily make everything fine. (28a–b)

The visible world cannot be eternal, since it is subject to change, and nothing subject to change is eternal (28b–c): but in that case it must have come to be as a result of some cause (28c). And being 'the fairest of things generated', the visible world must have had 'the best of causes' (29a); consequently, its maker must have looked towards what is best, namely the eternal and unchanging, and modelled the world as far as possible in its likeness.

The idea that the world is fine rests ultimately on the notion that it exhibits structure and order, the attributes of Mind or intelligence. Anything created with intelligence is superior to anything without it; and for this reason, the Artisan made the world with Soul (30a–c; cf. *Philebus* 28d–e, 30a–e). Furthermore, Plato implicitly relies upon an application of P4 (§1*b* above), or PCS; whatever is responsible for something's being fine must itself be fine. But note that this now comes in two distinct forms, parallel to the two senses of 'from' isolated earlier (P5 and P6: §1*b*): both the maker and his models must be fine (in different ways, of course). If the Artisan is to make something fine, he must himself be fine (in the sense of disposed towards the best); equally he must employ something fine, in this case the eternal paradigms of the Forms. From these assumptions it follows, Timaeus says, that

141 according to the likely account we must say that this cosmos was generated as an ensouled intelligent animal, in reality through the providence of God. (*Timaeus* 30b–c)

Plato grounds his teleological conception of the universe in an argument from design:

(1) [AD] the world is such that it cannot have arisen except as a result of conscious creation;

(2) any such creator must be of supreme wisdom and goodness;

hence

(3) every facet of that creation must exhibit the goodness of the creator's design.

end p.110

[AD] is not as it stands conclusive. It requires supplemental assumptions, and (1) is obviously controversial. But neither is it negligible. We have already seen the difficulties Greek mechanist theoreticians were faced with in trying to account for the emergence of structure and organization in the world (I.4*b*–*d*), a problem perhaps particularly acute for the various varieties of atomism, the dominant ancient materialism (VI.1*d*, 2*a*–*e*). This is emphasized elsewhere:

142 socrates . Should we say, Protarchus, that everything, and what we call this Whole, is ruled by the power of irrationality and randomness, or the opposite, as those before us held, that Mind and a remarkable ordering fore- thought control it?
protarchus . There's no comparison, Socrates. What you have just now said seems blasphemous to me; while to say that Mind orders everything is adequate to the visual evidence of the universe, the sun, the moon, the stars, and all their revolutions. (Plato, *Philebus* 28d–e)

Lacking a mechanics that will account for the apparently uniform behaviour of the heavenly bodies, and in default (for the biosphere) of any concept of how cybernetic systems can become self-regulating and mimic purposive behaviour, the option of choosing a teleology invoking an actual creator (or, as I shall henceforth label it, a ‘directed teleology’)[38] becomes highly attractive, making premiss (1) at least a reasonable assumption. But that is precisely the status Timaeus claims for his story.

Timaeus next infers that the world, being good, must be unique, since God's purpose was to create the best world available, and one world among many could not be so pre-eminent (*Timaeus* 30d–31a; cf. 55c–d). This argument resembles that of Xenophanes for the uniqueness of God (I.3*a*); and it seems subject to similar objections. Why could God not have created indefinitely many worlds, each the best they could be? Plato's answer is that each would then be parts of a larger whole (i.e. the totality of worlds); but that larger whole would not itself be an organic unity: and, unity being better than diversity, the whole would not be as good as it might be (31a).

(b) The Construction of the Cosmos

Having established the teleological foundation for the explanation of the world, Timaeus proceeds to deduce its structure on the basis of the assumption that it is the work of reason. The world is created.[39] But anything created must be visible and tangible: it cannot be the former

unless it contains fire, or the latter unless it contains earth; and so the Artisan was constrained to produce both fire and earth (*Timaeus* 31b).

Plato further attempts to give an explanation for the interpolation of water and air between them, although it has to be said that the account is somewhat strained and smacks of Pythagorean number-mysticism (cf. I.4*a*). He puts the elements in geometrical proportion to one another, such that E:W::W:A::A:F; presumably these proportions are meant to represent somehow their relative densities. But the only ‘explanation’ for there being two intermediate terms is that the world is three-dimensional:

143 if the body of the universe had been a plane surface having no depth, one mean would have sufficed to bind itself with the others, but in fact it has to be solid, and solids are always connected with two middle terms. (32a–b)

No real reason is given *why* the world should be three-dimensional (except perhaps that otherwise it could not be visible or tangible). Aristotle (*On the Heavens* 1. 1. 268^{a}7–20) argues that the universe must be three- dimensional, since three is the maximum number of available dimensions; and some contemporary cosmologists hold, on the basis of (one version of) the ‘anthropic principle’, that there are reasons why the number of experienced spatial dimensions must be three.[40] None the less, Plato makes his universe spherical, on the grounds that ‘uniformity is incalculably better than its opposite’ (*Timaeus* 33b); and one would have thought that the more dimensions uniformity can be expressed in the better.

Even so, that does not account for the geometrical proportions. It was known in Plato's day that the classical geometrical problem of doubling the cube (i.e. finding the constant *c* such that a cube of side *c*. *l* would be double the volume of a cube with side *l*) reduced to finding two geometrical means between 1 and 2 (i.e. *n* and *m* such that 1: *n*:: *n*: *m*::

m:2): *n* then is the constant *c*. But that is a highly impressionistic way to relate three-dimensionality to two intermediate means.

Nevertheless, Plato tries as far as possible to generate the world on the basis of rational considerations (sometimes of an abstract and aesthetic nature) of what is for the best. Thus the creator makes use of all the available material (since it is better to let none go to waste: 32c–33b), in the best possible spherical shape. Furthermore, although an animal, since it contains all there is, it has no need of ingestion and excretion like the living creatures which are its component parts, since 'the creator considered it

end p.112

better for it to be self-sufficient than for it to be reliant on other things (33d)'. Finally it moves with that motion best suited to intelligence, namely uniform circular motion (34a). Plato next has the Demiurge construct the heavenly bodies from two obliquely angled motions, of the Same (accounting for the diurnal rotation of the celestial sphere) and the Different (accounting for the movement of the bodies of the solar system along the ecliptic).[41] The construction is again obscurely geometrical, making use of the ratios obtaining between tones on the musical scale; Plato is impressed presumably for the same reasons as the Pythagoreans with the mathematical nature of harmony (34b–37c: cf. I.4*a*). The actual content of this matters less than its expression of Plato's belief that the universe is susceptible of a comprehensive account which will exhibit its overall rationality, even if the periodicities of the planets (upon which Plato thinks time itself depends) are extremely difficult to calculate, surpassing most people's understanding in their multiform intricacy (39c–d: see further §5 below).

The Same and the Different also figure, somewhat confusingly, in the construction of the human soul (41a–44d): once again subordinate creatures aspire as far as possible to the perfection of the highest beings. Plato next (45b–46d) offers a detailed account of the physiology of vision, explaining (at least to his own satisfaction) the mechanisms by which we can see the world around us. But he concludes his physiological explanations as follows:

144 all these are co-operative causes (*sunaitia*)[42] which God uses as his ministers in bringing to fulfilment the idea of the best as far as it is possible. These are, however, thought by most people not to be co-operative causes but causes . . . but these cannot have any reason (*logos*) or Mind (*nous*) in them. The lover of *nous* and of understanding (*epistēmē*) should first of all pursue the causes of intelligent nature, and only afterwards those which are such as to move other things of necessity because they themselves are moved . . . both types of causes should be enumerated, but those which are intelligent creators of fine and good things should be differentiated from those which lack Mind and produce haphazard, disorderly results. Let these things be said

end p.113

concerning the accessory causes (*summetaitia*)[43] of the eyes possessing the capacity (*dunamis*) which they now have. (46c–e)

This mirrors the *Phaedo* account (§1*a*, **120–3**). The *sunaitia* (cf. X.1*b*, **353**; VII.1*c*, **257**) may be compared with the prerequisites, the things 'without which' of **124** (VII.1*e*, **263**; X.2*c*, **364**; XI.3*a*, **401**; XI.3*b*, **402**; XI,3*c*, **404**), although the comparison is not exact (the prerequisites are not themselves causes at all, whereas here the *sunaitia* actually do something). Only when tied down with intelligence and purpose will the material factors produce order rather than 'haphazard, disorderly results'. Mind is needed to inject direction and structure into the system, while the fact that the world exhibits such structure, while explanations which appeal to the Good and the Best can be seen to be fruitful in particular instances, underwrites the legitimacy of extending the teleological principle (on the basis of argument AD: § *a* above) to more recalcitrant cases.

The teleological hypothesis thus becomes a regulative principle on explanatory adequacy. So, in the case of the eyes, we must explain not merely how they work, but also why it is a good thing that they do, in terms of the benefits that accrue to the sighted. Plato's answer is that without them we should never have been able to observe the eternal motions of the heavenly bodies and from them derive the notions of number, time, and ultimately philosophy, 'the greatest of gods that has or will come from gods to mortals' (*Timaeus* 47a–b).[44] Once again, material explanations are not so much false as inadequate.

Material facts, then, constrain what the creator may accomplish by restricting the range of causal possibilities from which he may choose. But they do not determine that choice. For this reason, material alone can never supply a full explanation consistent with P1 and P2. Even so, these constraints are important. They will explain, for instance, why the creator could not make human beings both intelligent and long-lived—bone, the enduring substance, cannot for material reasons bear Mind, while the soft tissues which can are, equally for reasons of material necessity, subject to inevitable decay. The creator is forced to decide between two causally

possible alternatives, neither of which is the best logically possible case (75a–c). Yet, within the limits thus imposed, the choice can still be seen to be good (better to be intelligently ephemeral than durably dull).

For this reason the discussion changes course at 48e:

145 virtually all of what has gone before has shown how things were crafted according to Mind—but we must also add to the account the results of necessity. For the generation of the cosmos came to be out of a mixture and composition of necessity and mind. Mind controlled necessity by persuading it for the most part[45] to direct the outcomes for the best, and in this way, with necessity subordinated to reasonable persuasion thus at the outset, this Whole was created. So if anyone wants to say truly how it came to be with these things in mind, the wandering cause (*planōmenē aitia*) must be mixed into it. (47e–48a)

The 'wandering cause' is material necessity, 'wandering' precisely because it does not on its own determine stable outcomes.[46] So, Timaeus says, they need to start once more at the beginning, this time concentrating on the original nature of earth, water, air, and fire, and their attributes (48b). Once more the principle of likelihood is to be adhered to: such an account can deliver nothing more. The investigation is to be, in a sense, a priori: we are to determine what the basic metaphysical structure of the elements must be on the basis of how they are required to interact in order to produce the visible world; but such an investigation will yield results that are at best probable.

(c) The Receptacle

146 Our starting-point in regard to the Whole must be more fully divided than before: then we distinguished two forms, while now we need to distinguish a third and distinct class. (*Timaeus* 48e)

end p.115

So Timaeus begins his second cosmological treatment. The 'two forms' of the previous story are the eternal Forms the Artisan copied, and their sensible representations. But this will no longer suffice: and we need a third category, even if

147it is a form difficult and obscure to express in words. What power (*dunamis*) and nature (*phusis*) must it be supposed to have? This most particularly: it is the receptacle (*hupodochē*) and as it were the nurse of all generation. (49a)

People wrongly take earth, water, air, and fire to be basic elements, like the letters of the alphabet: but in reality they are more like syllables (48b–c). The difficulty Timaeus confronts is occasioned by the intertransmutability of the elements. If the elements themselves can intertransmute, what underlies those elemental transformations (49c)?

148 So since none of them ever appears the same, how could one fail to shrink from maintaining that any one is certainly one rather than another? They are not; but it will be by far the safest for us to make the following suppositions and speak as follows: whenever we see something that is always in the process of changing into something else, like fire, we should not call it a 'this' but a 'such and such'; and similarly water we should call not 'this' but 'such and such.' (49d)

The point of this linguistic stipulation is to avoid the suggestion, implicit in the use of the demonstrative, that what is being referred to is in some sense permanent: thus we may speak of something's being fiery, but not of its being fire.[47]

That may seem trivial—but an important point lurks behind it. A Heraclitean flux threatens to undermine the referential function of language, and with it language itself (as Plato argues at *Theaetetus* 179d–183b).[48] Moreover, such a position furthers the distinction between object and property: we are now being invited to refer to things as bearers of properties, but not as though they are identical with those properties. At the most fundamental level, we are left with a pure property-bearer which in and of itself has no properties, something akin to the *prima materia* of medieval Aristotelianism.[49] Indeed, Plato's later assimilation of the

end p.116

receptacle to space (*Timaeus* 52a–b) appears to anticipate the later concept of matter as extension endowed with properties, although that anticipation may be misleading.[50]

At all events, the receptacle is neutral with regard to properties (50b–c; Plato compares it to the odourless base oil used by perfume manufacturers: 50e). Thus we have

149 three categories: that which is generated, that in which it is generated, and that in the likeness of which that which is generated is made. (50c–d)

Timaeus compares the paradigm to a father, the receptacle to a mother, and the thing generated to their offspring, which strongly suggests that he does wish to conceptualize the Forms as causes, although he allows that 'how this happens is wonderful and hard to describe' (50c–d).[51] That there are such originative forms, whose earthly copies exist when regions of the receptacle take on their essential characteristics, is argued at 50b–d: the possibility of knowledge, whose objects must be eternal (§3*b* above), guarantees the existence of the eternal exemplars.

Once this is established,

150 it must be agreed that there exists first of all the Form, ungenerated and indestructible, which admits of no alteration or combination, which is invisible and imperceptible to the other senses, which is apprehended by thought; second is that which is similar both in name and form to it, but is perceptible, generated, and ever-changing, which comes into being in a particular place, and passes out of existence there too, which is grasped by opinion along with perception; and there is a third category of eternal space, which does not admit of destruction, but which provides a seat for everything which is generated, and which is grasped without the senses, but by a sort of bastard reasoning. (51e–52b; cf. **449**)

Before the Artisan imposed order on the world, there was space filled with 'disorderly motions', the basic material of the universe with no structure to it (52d–e). Gradually, however, the elements within it are partially sifted out as a result of the internal dynamics of the system (Plato compares the receptacle to a sieve: 52e). Plato appears to envisage some sort of anti-entropic principle at work before the Artisan appears on the scene: even in this primitive, unelaborated stage, various parts of matter exhibit to a certain degree the attributes of earth, water, air, and fire (presumably under the influence of the Forms), but not yet to their fullest extent (53a–b).

(d) The Elements

It is not clear why Plato feels he must postulate this preliminary, pre- organizational stage of development: but one answer that suggests itself is this. The world, even after the intervention of the Artisan, must still roll forward under its own steam (Plato does not envisage the continual intervention of God in the world as a conserving cause); it must, then, have dynamism built into it, after the manner of Anaximander (I.2*d*, **7**(ii)) and Empedocles (I.4*b*). If that is so, we should expect these constant, disorderly motions periodically to produce areas of elemental differentiation within the mixture.

At all events, mere random action of the sort postulated by Democritus and the Atomists (I.4*d*; VI.1*d*) will not generate the cosmos as we know it (cf. *Laws* 10. 886d–e, 889a–e). God is needed for that: and the first thing he did was to settle upon some definite number for the second-order elements, and on some relations between them, according to what was best (*Timaeus* 53b–c; this recalls 31b–32b). Plato now proceeds to generate the four elements in a somewhat different manner, beginning from abstract geometry. There are, he claims, two basic types of right-angled triangle: (*a*) the right-angled isosceles, and (*b*) the right-angled triangle that results from the bisection of the equilateral triangle (53c–d).[52] These, it is claimed, are the basis of the elements 'following the necessary reasoning of probability' (53d; again, this is not explained). Plato next shows how four of the five regular solids may be generated from combinations of these triangles (the pyramid, octahedron, and icosahedron from those of type (*a*), the cube from type (*b*): 53d–54a). Transmutation of the elements, then, involves a rearrangement of the triangles. It follows immediately from this that earth, contrary to what has been implied before (49c), cannot intertransmute with the other elements (54c).

In what follows, Plato elaborates and clarifies the account: the cube is 'firmer', hence more appropriate for earth (55d–e); of the other elements the 'least mobile' (the icosahedron) is assigned to water, the most mobile (the pyramid) to fire, and the intermediate (the octahedron) to air (56a); furthermore, fire should have the sharpest angles (of the pyramid), since it is sharper (56a–b). These basic structural properties are supposed to account for the phenomenal properties of the elements (an account which is

worked out in more detail at 61c–64a), as well as their modes of interaction and the manner of their intertransmutation (56c–57c: a unit of water, containing twenty faces or forty triangles, may thus divide into two eight- faced units of air and one pyramid of fire: 56d).

Plato's chemical scheme may be fanciful—but it represents a thorough- going attempt at reductive explanation (cf. Introduction, p. 3). In principle all the phenomenal properties of objects are to be explained in terms of the basic geometrical structures of their constituents: and at 64a–68d, Plato offers similar reductive accounts of pleasure and pain, the senses, and of colour phenomena. What matters is not so much that these explanations are empirically inadequate, but that they are attempted at all. Plato tries, in a more systematic and detailed manner than any of his predecessors (in particular Democritus, against whose conception of the world as a chance agglomeration of haphazardly interacting particles Plato is particularly arguing here: cf. VI.1*d*), to render the natural world susceptible of an orderly and precise explanation.[53]

This construction in some ways mirrors passages to be found in *Philebus*, and we may conclude by drawing attention briefly to them. *Philebus* 16c distinguishes between the unlimited (*to apeiron*) and the limit (*peras*) which may be imposed upon it.[54] The limits, or limiters, are numerical and quantitative in form. For example, the range of audible sound is unlimited (presumably in that it is a continuum, not that it may be arbitrarily high or low): musical theory consists in imposing order, in the form of mathematized intervals, on that indeterminate continuous range (17b–e). Plato's intent here is to say something about the nature of *technē*, organized skill: but his remarks carry over into the metaphysics of explanation. Essentially explanation (and with it the ability to learn and teach skills) involves the imposition of order on chaos. This point is applied a little later on to the construction of the cosmos. There are two basic principles: the unlimited (which may be assimilated to the receptacle of the *Timaeus*)[55] and the limit; from these there

arises a third category of things composed of the two (presumably here phenomenal objects,
end p.119

or at the very least phenomenal instantiations of properties: *Philebus* 25b–26d).
Finally, Plato posits a fourth category, that of the agent responsible for the imposition of limit on the indeterminate (26e–27c). These four categories, Plato suggests, are sufficient to generate the entire sensible world (30a–d). And the fact that Plato requires the fourth category shows the extent to which he is now able to distinguish between purely formal analyses of the structure of things (on the lines of the 'safe and simple-minded causes' of *Phaedo* 100d–e: **128**) and a causal account of how they come to be the way they are.

5. Coda: Mathematical and Astronomical Theory

The *Timaeus* offers an account of the structure of the cosmos, and it contains the elements of an astronomical theory which merits some brief consideration. Interest in formulating accurate calendars, and hence in discovering such quantities as the exact length of the solar year and the length of the seasons, emerges very early in the history of civilization in the West, in Babylonia and Egypt (and was central to the concerns of other quite independent civilizations such as those of ancient Meso-America). Such an interest is of course partly driven by the desire to place agriculture on a firm foundation: and the links between astronomical calculation and the proper scientific attitude to planting, sowing, and reaping is to be found throughout antiquity.
On the other hand, as the ancients well knew, agriculture is not such an exact science as to require a calendar accurate to fractions of a day in a year; and nor does the other central issue which obsessed the early calendar-theorists, namely the determination of the precise numerical relation between the solar year and the lunar month, seem to be of any immediately compelling practical importance.[56] In fact, a desire to determine the precise dates for the appropriate performance of seasonal rituals seems to have provided the principal impetus for the development of observational astronomy, including the recognition and naming of the various constellations, as a prerequisite for accurate calendar computation.[57]
end p.120

But whatever its motivations, the calculation of the periodicities of the lunar month and the solar year attained to a new precision in the Greek world towards the end of the fifth century bc , with the work of Meton, who postulated a cycle in which nineteen solar years were correlated with 235 lunar months. Subsequently these figures were refined and incorporated into cycles of different lengths, as the Greeks produced closer and closer approximations to the lengths of the lunar month and the mean solar year.[58] At roughly the same time Euctemon recognized the inequality of the seasons, although the relative lengths he ascribed to them were inaccurate.[59]

Plato admired mathematics, and encouraged its practice in the Academy (although it is not clear how much of his colleagues' advanced researches he genuinely understood). In the *Republic*, mathematical studies are central to the curriculum prescribed for the would-be philosopher-king (*Republic* 7. 522c–528e), and one of the key branches of mathematics is astronomy. But when Glaucon applauds Socrates' prescription because of the practical advantages it affords the farmer, navigator, and general (7. 527d), Socrates takes him to task (as he had done in the case of geometry) for missing the real point: the visible firmament is but a pale shadow of the genuine, unchanging reality towards which the inward gaze of the philosopher should be turned (7. 528e–530b). In fact, observational astronomy is to be abandoned in favour of a theoretical astronomy construed on the lines of abstract geometry (7. 530b–c).

These remarks have frequently been construed as indicating Plato's dismissive attitude towards observational astronomy, indeed towards things in the visible realm as a whole: but this has undoubtedly been exaggerated.[60] Rather, it seems that Plato sought to emphasize the importance of abstract theorizing not at the expense of observation, but in order to provide an explanatory framework within which to make sense of the increasing proliferation of observational data which was becoming available in the period. In his last work, the *Laws*, astronomy is equally central to the educational system:

151 athenian . Next look at astronomical knowledge and whether we should teach it to children or not.

cleinias . Go ahead.

ath . Well, there is something absolutely remarkable about this, something indeed completely unacceptable.

cl . What?

ath . We say that we should not investigate the greatest god and whole, the cosmos, nor busy ourselves in seeking its causes, since it is not holy to do so. But it seems that doing the precise opposite would be the right thing to do. . . . Whenever anyone believes that a science is fine and true, beneficial to the city, and altogether dear to God, in no way can he refrain from speaking of it.

cl . What you say is plausible: but what such science shall we discover of the stars?

ath . My friends, at the moment virtually all the Greeks are wrong about those great gods, the Sun and the Moon.

cl . Wrong in what way?

ath . We say that they, and some other stars with them, never travel the same road, and call them planets. (Plato, *Laws* 7. 820e–821b; 12. 966d–968a)

The falsehood is not the suggestion that the planets' motions are in some sense irregular, both in that their periods do not synchronize either with each other or with the solar year and (in the case of the outer planets at least) that they exhibit station and retrogression, which were well known in Plato's time. Rather what is false is the view that they exhibit no order or regularity whatsoever, that their motions are entirely random. Indeed it is said that Plato set it as a problem to be worked upon in his Academy to produce an account of these apparently multifarious motions in terms of some small, orderly set of fundamental motions (Simplicius, *On 'On the Heavens'* 488: **197**), a task triumphantly undertaken by

Eudoxus and others (V.2*f*). Plato's own astronomical pronouncements are a good deal more metaphorical in tone (in particular those to be found on the Myth of Er at the end of the *Republic*: 10. 614c–617d); but they exhibit some relatively advanced mathematical understanding in addition to embodying the sound, reductive methodological precept just mentioned.

In the cosmological section of the *Timaeus*, Plato first describes, in obscurely mathematical terms, the composition of the elemental basics of sameness, difference, and being (*Timaeus* 34c–36b); then this mixture is fashioned by the Demiurge into a shape like a letter 'X', each of whose extremities is then bent round to form two circles bisecting one another at an angle (36b–c) representing that of the ecliptic. The outer 'circle of the same' revolves in one direction, the inner 'circle of the different' moves in the opposite direction, and is divided into seven (36c–d), three of which (i.e. the Sun, Mercury, and Venus) move at the same (mean) velocity, while the others (the Moon, Mars, Jupiter, and Saturn) move at different rates 'although proportionally' (36d).

end p.122

As we noted above (§4*b*), teleological concerns drive the construction, and impel Plato to impute a greater degree of mathematical harmony to the overall structure, in neo-Pythagorean manner, than that which actually obtains. None the less, the solar system he describes is a moderately accurate one. A passage a page or so later confirms this impression, as well as being representative of his concerns:

152 the Sun, the Moon, and the five other stars which are called planets, were generated to differentiate and preserve the numbers of time; and God, after making each of their bodies, placed them in orbits which the circle of the Different moved, seven orbits for seven stars, the Moon in the first orbit around the Earth, the Sun in the second above the Earth, then the Morning Star [i.e. Venus] and that said to be sacred to Hermes [i.e. Mercury], moving in circles at the same rate as the Sun, although having a share of contrary power,[61] for which reason the Sun and Mercury and Venus in the same way overtake and are overtaken by each other. (38c–d)

All of this was done in order to mark time, first of all by the fundamental measure of night and day, and then by the lunar month and the solar year (39c):

153 however men, with a very few exceptions, are ignorant of the other periods, and neither name them nor measure them numerically against each other, and so they are more or less ignorant that time consists in their wanderings, being intractable in number and marvellously intricate. (39c–d)

The detailed working-out of the picture, then, was left to others, preeminently Eudoxus. But Plato is in no doubt either about the basic structure he wishes to ascribe to the heavens (which achieves a tolerable degree of observational understanding: he was apparently also aware of the stations and retrogressions of the outer planets: 39d), nor of the general metaphysical, teleological morals he wishes to draw therefrom. The heavenly bodies must be viewed as ensouled and divine, since in no other way can their perfection and regularity of movement be explained (*Laws* 12. 966d–967d); and the divinity and rationality of the heavens is the subject of particular emphasis in *Epinomis*. In this possibly spurious[62] dialogue, Plato (or one of his successors) repeats the claim that the

orderliness of the heavens demands that they be considered as divine (*Epinomis* 982a–983c),
end p.123

as they dance their eternal dance (982d–e; cf. *Timaeus* 40c) under the influence of their eight distinct cosmic powers (*Epinomis* 986a–988e).

6. Conclusions

With Plato we discover the first (surviving) genuinely philosophical reflections on the nature of cause and explanation, and their relations with other fundamental metaphysical concepts. In the *Phaedo*, he signals his dissatisfaction with prevailing materialist conceptions of physical explanation (§1*a*), and sketches an account in which invariable causal properties are to be invoked, along with a teleological appeal to the good, in order to ground explanations, an investigation which leads him to postulate a series of conditions which causes must meet (P1–P8, A1–A2: §1*b*).

He adapts Alcmaeon's immortality argument (§2*a*; cf. I.3*c*), and applies similar conditions of causal intelligibility in the final argument for the immortality of the soul in the *Phaedo* (§2*b*). Those arguments make use of the notion of an *archē*, or ultimate principle, a notion further explored in connection with the concept of hypothetical investigation (§§3*a–b*); finally, the concept of knowledge itself is firmly linked to that of causal understanding: to know something is to know why it is the case (§3*c*).

These concerns all come together in the *Timaeus* account of the generation of the cosmos: a divine artificer created the universe and its components with an eye to what was best, fashioning as far as material exigencies allow material representations of pure formal realities (§§4*a–b*); in the course of so doing, he elaborates a notion of a material substratum, itself without properties but fit for the reception of them, which was to be of enormous metaphysical influence (§4*c*), and sketches an account of chemical interaction (§4*d*). Plato's self-conscious concern with the proper form of explanation, his insistence on the necessity of teleology, and his account of the role of soul in the overall functioning of the universe, all attest to his concern with exhibiting the world and its contents as being fundamentally intelligible.
end p.124

IV Aristotle: Explanation and Nature

R. J. Hankinson

1. The Structure of Reality

(a) The Importance of Nature

Plato's world of eternal and unchanging Forms, imperfectly represented in matter by a divine Artisan, contrasts sharply with the various mechanistic *Weltanschauungen*, of which atomism was, by the fourth century at least, the most prominent (VI.1). This debate was to persist throughout the ancient world. Atomistic mechanism got a shot in the arm from Epicurus (VI.2), while the Stoics adopted a divine teleology (VII.3*a*; cf. XI.3*a*). The choice seems simple: either show how a structured, regular world could arise out of undirected processes, or inject intelligence into the system. This was how Aristotle (384–322 bc), when still a young acolyte of Plato, saw matters. Cicero (*On the Nature of the Gods* 2. 95 = Fr. 12) preserves Aristotle's own cave-image: if troglodytes were brought on a sudden into the upper world, they would immediately suppose it to have been intelligently arranged. But Aristotle grew to abandon this view; although he believes in a divine being, the Prime Mover is not the efficient cause of action in the Universe, and plays no part in constructing or arranging it (§3*d*, 4*a* below; V.2*g*).

But, although he rejects the divine Artificer, Aristotle does not resort to a pure mechanism of random forces. Instead he seeks to find a middle way between the two positions, one which relies heavily on the notion of Nature, or *phusis*. *Phusis* is a many-faceted concept in Aristotle (a chapter of his philosophical lexicon *Metaphysics* 5. 4 is devoted to disentangling its complexities). But broadly we may distinguish two principal senses: (*a*) Nature as a whole (i.e. the totality of the natural world; (*b*) 'the essence of those things which have an internal principle of change' which is 'the primary and proper sense of *phusis*' (*Metaphysics* 5. 4. 1015^{a}13–14). The world is divided into things which exist by nature, and those which exist 'for other reasons' (*Physics* 2. 1. 192^{b}8–9; cf. *Metaphysics* 7. 9).

The paradigm cases of the former are living things: plants, animals, and the heavenly bodies; but their elements, earth, water, air, and fire are also natural. The latter are represented by artefacts like beds. The basis of this distinction is a causal one. Natural objects possess their own internal

principles (Aristotle uses the word *archē*) of growth, movement, and rest (*Physics* 2. 1. 192^{b}13–14). By contrast, objects that do not exist by nature have no such internal principles; or rather they do not have them *qua* the thing that they are, but only derivatively, via the elements from which they are constructed.

Aristotle pointedly retails an example owed to an Antiphon: if we planted a bed, it will not procreate baby beds (if anything, it would propagate a tree: ibid. 2. 1. 193^{a}9–17). It is not a bed by nature, but rather because an artisan has imposed this structure onto the material. Natural objects require no artificers, since the active and the passive coincide in them (ibid. 2. 1. 192^{b}23–7; cf. 2. 8. 199^{b}30–2), and their internal structures are such that, material constraints willing, they will of their own devices grow into mature instances of their species, and reproduce. This goal- directedness, then, is an intrinsic feature of natural objects, which seek (metaphorically) to realize their natures; it requires no external intelligent imposition.

This distinction is born out of a desire to explain, which Aristotle takes to characterize humanity: 'all human beings desire by nature to know' (*Metaphysics* 1. 1. 980^{a}21). But knowledge, properly so-called, consists in specifying the reasons why things are the way they are. In the case of natural objects, then, knowledge consists in understanding their natures, or, in a phrase he employs widely, in grasping the what-it-is-to-be of those things. As the natures of things are intrinsic to them, Aristotle has no need at this point to rely upon a divine creator. Herein lies Aristotle's great departure from his Platonic inheritance. While Plato extends the model of the craftsman and his materials to account for the structure and organization of the perceptible world (III.4*a*), the mature Aristotle sharply distinguishes the two. Properly understood, form is only *imposed* upon matter in the case of artefacts; in natural objects, form is internal to the objects themselves (except at the moment of generation: §2*a* below).

Beds have no natures, since they neither grow, reproduce, nor locomote (all of which are properties Aristotle ascribes to different levels of the soul, the first two common to plants and animals, the last belonging to animals alone: *On the Soul* 2. 1–4; §4*a* below). The elements neither grow nor reproduce; but they have natures in so far as they have innate tendencies to move towards determinate regions of the Universe, and having got there to stay there (*On the Heavens* 1. 2–4, 4. 1–6). Thus whatever a bed does by nature (falls, floats, burns, and so on) can be explained at the level of the properties of its material constituents. These are real, natural properties, but not properties it has in virtue of its being a bed. We may speak loosely of the form of the bed—but it is not, for Aristotle, real form at all. Rather it is a superficial set of properties temporarily induced in the material.

This conception of something's nature as an internal principle of change is thus absolutely fundamental to Aristotle's explanatory scheme of things. That it is internal means that there is in a sense no distinction between the cause of a natural object's growth and the object itself: what is to be an *F* (where *F* picks out some natural kind) is to have the internal nisus for the realization of *F*-ness in its fullest form. This conception is teleological, but the teleology is immanent in natural things, Aristotelian substances (*ousiai*), not imposed by a designing Mind.

But things are not uniformly successful in their drive to achieve perfection. They may not survive infancy; but even if they do, they may imperfectly instantiate the kinds of which they are members. They may be deformed, or deficient, in numerous ways. Thus to be possessed of form, in the sense of having such an internal drive, is not a sufficient condition of its being realized: material factors may impede or prevent it altogether. But in that case, Aristotle thinks, there should always be an explanation available as to why the substance in question failed to achieve full maturity. Aristotle conceives of form as being an inner dynamic potentiality struggling against material constraints for its consummation in resistant and recalcitrant matter.

The concept of potentiality (and its relative actuality) is central to Aristotelian metaphysics (see *Metaphysics* 9). For an object *a* to be potentially *F*, more is required than that its being *F* is logically (or indeed causally) possible: it must be written into *a*'s structure that, other things being equal, it *will* be *F*. While the pile of wood in my backyard has the causal potentiality for being a table (if I get around to building one, and if I possess the requisite skills), that potentiality is purely passive in nature: the wood has no active internal nisus towards tablehood. By contrast, the acorn really is potentially an oak: it requires no external artisan, and, if the conditions are favourable, will become one as a result of its own internal structural processes.

But to say that it will, other things being equal, become an oak tree requires careful handling. Aristotle was perfectly well aware of nature's prodigality (cf. *Generation of Animals* 3. 1. 751^a25–30, on the number of fish-eggs). Thus, for *a* strictly to be potentially *F*, it is not necessary that most *a*-type things do as a matter of fact become *F*.[1] Rather if conditions are right, *a* will become *F* under its own steam, as it were. Conversely, if most *a*s do become *F* (where *a*s form a properly delimited natural class), then we may infer it is natural for *a*s to develop in that way.
end p.127

(b) Nature and Change

At the beginning of *Physics* 2, Aristotle wrestles with the concept of nature. Does it signify the matter out of which things are made, or does it rather pick out their form (*Physics* 2. 1. 192^b32–193^b21)? In fact it does both (2. 2. 194^a12–27; cf. *Metaphysics* 7. 1–3). Moreover, physical explanation involves ends or purposes (*Physics* 2. 2. 194^a27–194^b9). Nature is in a sense analogous with art (Aristotle exploits this analogy elsewhere: *Metaphysics* 7. 7–9), and the artist selects and deploys his material with a view to some particular end (thus the helmsman directs the shipbuilder by giving him the specifications with which to construct a rudder, while the latter knows how to realize those specifications in the material: *Physics* 2. 2. 194^b5–7):

154 thus in the case of things done by art (*technē*)[2] we produce the material for the sake of the function (*ergon*), while in natural cases it exists already. (ibid. 2. 2. 194^b7–9)

Aristotle formally analyses generation in such a way as to defuse Parmenides' objections to it on the basis of the distinction between the causal and the material senses of 'from':

155 something must always underlie what comes to be, and this, while numerically one, is not one in form (by 'form' and 'account' I mean the same thing): for being a man and being uncultured are not the same, and the one persists while the other does not; that which is not an opposite persists (for the man persists), while the cultured and the uncultured do not (and nor does the combination of the two, e.g. 'uncultured man'). We generally speak of something coming to be from something (and not of it coming to be some particular thing) in respect of the non-persisting things (e.g. the cultured comes to be from the uncultured, and not from the man). Even so, we sometimes speak thus in respect of what persists; for we say that the sculpture comes to be from the bronze, not that the bronze comes to be a sculpture. However, we speak of coming to be from the opposite and the non-persisting in both ways, this coming to be from that, and this coming to be that: for we say both that the cultured comes to be from the uncultured and that the uncultured man comes to be cultured. (ibid. 1. 7. 190^a14–31; cf. *Metaphysics* 5. 24; *Generation of Animals* 1. 18. 724^a20–34; cf. III.1*b*: P5 and P6).

The standard form of change is

(C) *x* comes to be *F* from having been not-*F* (cf. A1 and A2: III.1*b*),
end p.128

in which a property *F* is at one time predicated of a persisting subject *x* and at another time is not. Thus in all cases, something must persist through the change. However, Aristotle further notes an ambiguity in the concept of coming to be: either (*a*) something may come to be *simpliciter*, i.e. come into existence—or (*b*) it may come to be *F* for some value of *F* (*Physics* 1. 7. 190^{a}31–3). C clearly models case (*b*), since changes in categories such as quality, quantity, relation, and so on are predicated of subjects. However, Aristotle holds, this analysis can be extended to (*a*)-cases as well, given that there is no such thing as creation *ex nihilo*:

156 that the substances themselves and everything else which exists *simpliciter*[3] come to be from some underlying substrate (*hupokeimenon*) will become apparent to one who investigates it: for there is always something which already exists out of which what is generated comes to be, for example plants and animals out of seed. (ibid. 1. 7. 190^{b}1–5)

Thus, in the case of substantial change, '*x*' refers to a parcel of matter, rather than an individual substance, while '*F*' will designate a substance- kind (i.e. 'man' or 'horse'). All change, then, presupposes that something, the substrate, persists through the change, as well as something in respect of which the change takes place:

157 consequently, it is clear from what has been said that everything which comes to be is always composite, and that there is something which comes to be, and that which comes to be this, and this latter is twofold, since it can refer both to the substrate and to the opposite. By opposite I mean e.g. 'uncultured', by substrate the individual man. (ibid. 1. 7. 190^{b}11–14)

Properly speaking, then, any analysis of change will have three components: the prior 'opposite' state (which Aristotle calls 'privation'), the state of being not-*F* (or at any rate being less *F*: in *Physics* 5. 1, Aristotle describes such intermediate changes in terms of change between polar opposites); the form which is generated in the change; and the material substrate, or basis for the change. Moreover, Aristotle accepts, at least in some circumstances, PCS (I.3*d*; cf. P4: III.1*b*): whatever it is that makes *x F* must itself be *F*, indeed paradigmatically whatever is the genuine cause of *F*-ness will be that which is *F* to the highest degree (the terminology of PCS is ultimately owed to Aristotle too: *Physics* 8. 5. 257^{b}7–14). Thus 'fire, which is the cause of heat in everything hot, is the hottest of things' (*Metaphysics* 2. 1. 993^{b}24–6; cf. *Posterior Analytics* 1. 2. 72^{a}29–30). Here Aristotle welcomes, and expands upon, some of his Platonic heritage (III.1*b*; cf. XII.2*b*, 3*b*), in the course of putting to rest Parmenidean qualms about motion (I.3*d*).
end p.129

Aristotelian matter is not simply stuff. Rather the physical Universe is a hierarchy of matter and form, at the most fundamental level of which exist the elements earth, water, air, and fire (and the element of the heavenly bodies, the ether). The substances (animals and plants) that are the basic ontological components of Aristotle's world are ultimately composed of these elements, at several removes of elaboration. Each stage upwards from the fundamental constituents involves a conjunction of matter and form. Matter is, for Aristotle, a relational concept (*Physics* 2. 2. 194^{b}8–9), and there is no such thing as matter as such: matter is the matter *for* something. And, crudely, the matter for any particular object *considered at a certain level in the hierarchy* is what exists at the next level down in the structure. Thus the elements are the matter for uniform[4] substances (such as blood or bone), which are themselves the matter for the organic parts, which are in turn matter for the animal. What counts as the matter depends precisely on the level at which you pick the object out.

However, Aristotle's metaphysics demands that there is (at least in the case of properly formed individuals) an appropriate level at which to pick things out: thus while it is true to say of Socrates that he is this collection of blood, bones, and flesh, none the less that designation does not say *what* he is, namely a human being. You pick an individual out correctly just in case you designate it at the highest level of form to which it attains. What *this* is (where ‘this’ points to Socrates) is a human being, not simply his set of component parts: for what it is to be Socrates is for these parts to be organized in characteristically human fashion.[5] And compatibly with this, what it is to come to be a human being is for these material parts to be suitably functionally organized as a whole. But the material (in a sense at least) precedes that of which it is the matter (*Metaphysics* 7. 7. $1032^{b}30$–$1033^{a}2$); and consequently Aristotle too rejects generation *ex nihilo*.

Two final pieces are required to complete the metaphysical picture. Aristotle distinguishes essential from accidental properties; but Aristotelian essences are not just those properties an object cannot fail to have and still be that object. Some properties which belong universally and necessarily are not essential, for instance the angle-sum property of triangles (*Posterior Analytics* 2. 3. $90^{b}7$–14; cf. *Topics* 1. 5. $102^{a}18$–30). Essences are what is expressed in the definitions of things (*Posterior Analytics* 2. 3. $90^{b}30$–$91^{a}1$); and definitions for Aristotle are Lockean real definitions (cf. Locke 1690: 3. 3. 15–19), expressing things's real natures. Essence and nature are intimately linked, and the properties that make up the essence of something (and hence which figure in its definition) are

end p.130

supposed to explain and underlie all the object's other necessary attributes. Thus having angle-sum 2R *follows from* what it is to be a triangle, although it is not itself part of what it is to be a triangle; and if you want to know why every triangle has 2R, you need to turn to the definition of triangularity which expresses its essence.

But the relations that hold between essences and properties dependent upon them are not exclusively logical or formal. In the case of biological substances (as opposed to mathematical objects) the dependence is causal. At least as far as the natural world is concerned, some of a species' natural properties will none the less not be universally instantiated in the species. Not all men are bearded (*Posterior Analytics* 2. 12. $96^{a}8$–11; cf. *History of Animals* 3. 11. $518^{a}19$–$^{b}4$): in some cases material factors intervene to prevent the form of the creature from fully realizing itself.

Finally, we need briefly to examine the connection between the notions of potentiality and form. At conception, the embryo contains within itself the potentiality for growth into a complete animal; that is, an embryo *F* has the form of *F* at least potentially (*Generation of Animals* 2. 3. $737^{a}17$–34). Form-potentials are causal potencies in the objects themselves, which can be distinguished according to the hierarchical scheme outlined above. Thus the complex of matter that makes up the embryo has the potentiality to grow into a complete animal in virtue of its possessing (potentially) that animal's form;

it also has a number of other potentialities, but not *qua* the animal-kind in question. The acorn is potentially an oak because of its oak- form; in a sense, it is also potentially a bed, but not *qua* form of oak. It has that potentiality in a secondary manner (since if it becomes an oak it will of necessity be wooden, and wood can be fashioned into beds). Equally, it is edible (at least if you're a squirrel), but again not *qua* oak but rather *qua* the structure of its matter (it is vegetable). Thus potentialities divide in the same way as the matter–form hierarchy.

Appeal to potentialities is not, as it might first appear, simply explanatory Molière-ism (cf. III.1*b* n. 8); Aristotle does not merely redescribe effects and claim thereby to have isolated their causes. Rather he posits, as a causal hypothesis, the existence of a structure to account for the causal regularities exhibited in nature, from which we may advance to richer and more satisfying forms of explanation (this is part and parcel of his attempt to chart a middle course between a universe of random events on the one hand and one under direct divine control on the other: § *a* above, §3 below). Actuality and potentiality relate to matter and form (although their relations are not quite as tight as is sometimes supposed), since the proximate matter for something is that thing potentially: blood, bones, and flesh are the matter for, and are potentially, a human being. By contrast, actuality corresponds to fully realized form: to be an actual oak tree is just

to have brought to full fruition the form of the oak inherent potentially in the acorn.

2. Causation and Coincidence

(a) The Four Causes

The translation 'cause' for Aristotle's *aitia* is in some respects misleading,[6] but it is not disastrous: the explanatory categories which Aristotle isolates are clearly designed to enable us to give causal accounts of things. His principal concern is with generation; and that is evidently a causal concern. Moreover, his interests are not primarily linguistic—rather he wants to understand how things actually are. Proper knowledge, or understanding (*epistēmē*), involves knowing the fundamental structure of things. To this end, Aristotle distinguishes four general classes of explanation:

158(i) in one way the cause is said to be the existing thing out of which something comes to be, e.g. the bronze of the statue, or the silver of the phial, and their genera. (ii) Another is the form or the template (*paradeigma*): this is the formula (*logos*) of the what-it-is-to-be, and its genera. . . . (iii) Further more, that from which the primary origin (*archē*) of change and rest, e.g. the responsible deliberator, or the father of the child, and in general the agent of the thing produced, and the changer of the thing changed. (iv) Moreover, there is the end (*telos*). This is that for the sake of which, e.g. health of walking; for why does he walk? In order, we say, to be healthy, and in so saying we think that we have given the cause. (Aristotle, *Physics* 2. 3. 194^{b}23–35) (i–iv) describe respectively the material, formal, efficient, and final causes. You can pick out the explanans in different ways, for instance by referring to the statue's material as bronze, or generically as metal (cf. 2. 3. 195^{a}29–31: we may describe the cause of health as a doctor, or simply as someone skilful). Moreover, we may pick out something as *aitios* under a description which is incidental (*kata sumbebēkos*) to its responsibility, as for instance when we designate the cause of the sculpture as Polycleitus, or as ‘a man’, or even more generally ‘an animal’ (2. 3. 195^{a}33–^{b}1):

159 in the class of incidentals some are closer and some are further away [i.e. from the explanatorily appropriate designation], as for instance if ‘the pale and cultured individual’ was said to be responsible for the statue. (2. 3. 195^{b}1–3)

At least as far as agency is concerned, we can specify the cause as potential (‘the builder’) or actual (‘the builder actually building’: 2. 3.
end p.132

195^{b}4–6). Finally, we may treat the sculpture as a particular statue, or as a statue, or as a representation (2. 3. 195^{b}7–9). How we refer to the causal factors matters to explanation, which is intensional in form. But Aristotle's tolerance of generic or incidental designations in some contexts indicates his sensitivity to the fact that purely causal language is extensional,[7] and shows him to be concerned with cause as well as explanation. If my only concern is to designate the cause of some outcome, I need not worry whether I pick it out *qua* author of the result in question. I may refer to the sculptor as the man in the beret if I want to. But if I want to explain the result, I must refer to him as bearing the appropriate explanatory property (‘that sculptor’), or strictly speaking directly to the property (‘the art of sculpture’) itself (2. 3. 195^{a}33–^{b}3).

This suggests some points of contact between Aristotle's theory and that of Plato in the *Phaedo*. First of all, while Aristotle's examples of explanatory relations are categorially promiscuous (he includes agents and events as causes, events and objects as effects), he is primarily interested in the explanation of the generation of things; and he parses this, as Plato does, in terms of their coming to bear particular attributes (the special case of substantial generation involves the instantiation of a collection of essential attributes in some suitable material).

Thus what most properly explains the generation of a particular sculpture is the possession by the sculptor of the art of statuary. This to be construed as entailing that the agent or efficient cause (Polycleitus) has within himself the ability to produce the sculpture, and that involves his possessing the form which he is to induce in the material. The form itself is not generated (*Metaphysics* 7. 8. 1033^{a}24–^{b}20), since generation consists in the taking on of form, and so the form itself would have to be composed of matter and form: and so on, *ad infinitum*); rather the craftsman induces the form in the matter. What comes to be is a concrete particular (this sculpture; that brazen sphere); but both its matter and form pre-exist it (albeit in different senses).

Similarly for natural generation: the form of the human being pre-exists (in the semen) its enmatterment (in the menstrual fluid: *Generation of Animals* 1. 19. 727^{a}25–30, 2. 4. 739^{b}19–30); this is why human beings produce other human beings (*Metaphysics* 7. 8. 1033^{b}32). But crucially there is no need to invoke Platonic, hypostasized, metaphysically free-standing Forms: such things would either have to be inert, Aristotle thinks, and

end p.133

hence explanatorily useless, or else they would have to be acting all the time (*On Generation and Destruction* 2. 9. 335^{b}8–20). Only properties as possessed by actual individuals can be transmitted (this is one way in which actuality is prior to potentiality: §3*b* below; cf. V.2*g*, 198(ii–iii); Plotinus, XII.2*b*, 443), and so the Platonic entities are explanatorily redundant, even if they can be rendered intelligible.[8] In the case of artificial productions (the products of skill), the form is mediated by the mind (*Metaphysics* 7. 7. 1032^{b}1–25, 9. 1034^{a}24). The artisan works with a mental model of the finished product, and the mind's instruction enables that form to be transferred to some new material. Aristotle compares natural and artificial processes (*Metaphysics* 7. 8–9), but in contrast with Plato of the *Timaeus*, natural generation differs fundamentally from artificial production in that it requires no mental mediation: rather the transfer of form involved is direct. Consequently, artificial productions are only to a limited extent analogous to natural generation, since the mode of transfer of form is quite different in each case, although art imitates nature, and attempts to improve upon it, which, Aristotle thinks, entails that nature too be purposive (*Physics* 2. 8. 199^{a}15–21). Thus Aristotle frees biology from the shackles of direct intentionality.

The examples of formal causes in *Physics* 2. 3 tend to be mathematical: the ratio 2 : 1 is the formal cause of the octave (194^{b}27–9); similarly, being three-sided is the formal cause of a triangle's having angle-sum 2R. These mathematical formal causes are perhaps best treated as being part of the analysis of what it is to be the mathematical object in question; and they consequently hold invariably of whatever it is they are the causes of. However, Aristotle's interest in form is not exclusively or even predominantly connected with mathematics; and in the case of natural generation the form may fail of realization. In nature, then, unlike mathematics, the regularities which hold among objects and their features are not invariable. They are, however, regular: 'all things which come to be by nature do so either invariably or for the most part' (ibid. 2. 8. 198^{b}35–6: see § *b* below). Furthermore, while in the case of artefacts we can generally specify an end to their production distinct from their merely coming to exist (a sculptor might produce a bronze for the sake of profit, for instance), in the case of natural processes the goal, properly considered, simply *is* the realization of the form. An acorn has no purpose other than to become an oak (although Aristotle sometimes allows that some individual kinds may be for the sake of others: §3*d* below; compare the case of the organic

end p.134

parts: *Generation of Animals* 2. 6. 742^{a}16–^{b}17). It is in this sense that Aristotle claims that formal and final causes regularly coincide (at least in the case of naturally generated things: *Physics* 2. 7. 198^{a}25–8). The final cause or end to which the organism tends is the full realization of its latent form. Thus the formal cause is no mere analytical unpacking of what it is to be something—rather, at least as far as living things are concerned, it is the dynamic package of internal structural attributes in virtue of which the process of maturation unfolds (standardly at least) the way it does. But because in Aristotle's scheme of things the end attained is in a sense (although obviously not a temporal one:

Metaphysics 5. 11) prior to the process which brings it to fruition, it makes sense to appeal to the actualized end-state as part of the explanation of the process itself. This imports no intentionality, overt or otherwise. Rather Aristotle relies here on the perfectly ordinary intuition that it makes sense to say that the acorn *is growing into* an oak: the process is directed, and directed towards an end-state. It may not, for material reasons, attain it; but to describe it in any other way would be to assimilate it to mere random material interaction, which would be fundamentally to misunderstand the regular nature of the natural universe (see further §3*a* below).

(b) Chance and Spontaneity

Aristotle's explanatory categories are broad. Events may be the efficient causes of other events, both in particular cases (the Athenian raid on Sardis caused the Persian War: *Posterior Analytics* 2. 11. 94^{a}36–^{b}1; cf. *Physics* 2. 7. 198^{a}18), and in general, as the deprivation of light by the interposition of the earth is responsible for lunar eclipses (*Posterior Analytics* 2. 8. 93^{a}29–^{b}8). Facts may cause other facts, as the fact that planets are nearby explains why they shine without twinkling (ibid. 1. 13. 78^{a}29–^{b}4; cf. V.1*a*, **181**), and the deciduousness of broad-leaved trees is explained by their sap-coagulation (2. 16. 98^{a}35–^{b}38, 17. 99^{a}23–9: cf. V.1*a*, **180**). Actions, such as post-prandial perambulation, cause states, such as health (2. 11. 94^{b}8–26). Moreover, final causes may be physical states of maturity, or, in the case of human intentional action, desired outcomes (**158**(iv)); and the premisses of a syllogism may be considered as its material (*Physics* 2. 3. 195^{a}16). And while Aristotle tends to insist that, strictly speaking, causes and effects are coeval (when the builder stops building, the house stops coming to be: ibid. 195^{b}16–22), and that causes properly speaking are both necessary and sufficient for their effects (*Posterior Analytics* 2. 17), he will on occasion relax both strictures, and allow talk of non-sufficient causes, and causes preceding their effects in time (2. 12.
end p.135

95^{a}10–^{b}12), although in the latter case in particular his treatment is sketchy and unsatisfactory.[9]

But in spite of this catholicity of approach, not everything that occurs or comes to be can be given a complete explanation invoking all four causal factors. Some things have no real material causes; and some types of spontaneously generated animals have no efficient cause.[10] But more importantly, some lack final causes: sometimes things simply occur for no reason. Aristotle considers the correct analysis of these cases in *Physics* 2. 4–6.

He does so by introducing two terms, *automaton* and *tuchē*, neither of which goes comfortably into English. I shall translate them respectively as ‘chance’ and ‘luck’. The former is the wider, inclusive category: anything which happens for no purpose is *automaton*, but only those things which might have been (but as a matter of fact were not) done expressly for benefit can be classified under *tuchē* (ibid. 2. 6. 197^{a}36–7). Thus if you happen to meet someone who owes you money in the market-place, and recover the debt, the outcome is lucky (2. 5. 196^{b}34–197^{a}5, 197^{a}16), since you did not go there for that purpose although you might have done: 'to collect a debt' is a perfectly intelligible reason for doing so, although as it happens it was not on this occasion operative. Things occurring by chance include a horse's happening in its flight to reach a place of safety, and a tripod's happening to fall on its feet so someone can sit upon it (2. 6. 197^{b}14–19). Aristotle sums up:

160 consequently, it is clear that among things that come to be in a general manner for some purpose [sc. not with that purpose in mind], whenever something occurs such that its external cause did not bring it about for the sake of its outcome, we describe it as chance. While we ascribe to luck those random occurrences which are chosen [sc. although not chosen as such] by creatures which have the capacity for choice. (2. 6. 197^{b}19–22)

Aristotle is concerned only with events that might have been the results of purposive action. A chance event is not any random event: if a stone happens to fall from a roof and kills someone, that is *automaton*, since you can try to kill someone that way (2. 6. 197^{b}30–2); but a stone's simply falling to the ground does not even qualify as a chance outcome in this sense.

end p.136

Aristotle does not deny that there may be such random outcomes; but they are, from the point of view of Aristotelian science (i.e. the theory of the interaction of substances endowed with form), entirely uninteresting. Furthermore, Aristotle's account of chance is perfectly compatible with any version of determinism: chance events may be construed as the spatio- temporal intersection of causal chains, each of which is determinist in general character. In the case of the recovery of money, we may believe that the presence of both creditor and debtor in the market is explicable (Aristotle in fact says as much); and we may interpret that deterministically. What makes it a matter of chance is not that it might have come about otherwise (although nothing in what Aristotle says here commits him to determinism), but rather that there was nothing in the separate sequences of events that lead each to the market-place that had as its goal their coincidence there (2. 5. 197^{a}8–30; cf. **162**).[11]

Their chance conjunction is explanatorily underdetermined: there is no explanation in the strong sense (i.e. in terms of consciously entertained ends) for their both being there at the same time. By contrast, the natural processes of the world happen either always or for the most part (2. 8. 198^{b}34–5): and this means that they are susceptible of genuine final-cause explanations. But what does it mean to say that they happen either always or for the most part? Clearly there is no requirement that they be continuous (hot weather happens in the summer for the most part; but weather is not for the most part hot and summery: §3*d* below). The frequencies are relative frequencies: most men have beards (*Posterior Analytics* 2. 12. 96^{a}8–11), but most things are not bearded men.[12] Thus the claim that natural occurrences happen always or for the most part should be glossed something as follows:

(N) if q is natural, then given p, always or for the most part q,

where p isolates the relevant background domain in which we are interested (given that it is summer, hot, dry weather is to be expected). I formulate N in terms of propositional variables for generality's sake, in order to encompass lucky events, such as the meeting at the market-place; but in general Aristotle is concerned with the holding of properties by individuals, and hence

(N*) if it is natural for *F*s to be *G*s, then *F*s are always, or for the most part, *G*s

end p.137

may be closer to Aristotle's general position here. N* is expressed in a form appropriate to syllogistic inference (although the applicability of syllogistic to the logic of ‘for the most part’ is strictly limited).

These dyadic interpretations of ‘always or for the most part’ have the further advantage of stressing the causal nature of the connections involved. The whole point of the distinctions is precisely that something cannot be a matter of chance if it occurs regularly: and here of course ‘regularly’ does not mean ‘all or most of the time’ in any absolute sense. On the contrary, the occurrences of the conjunctions may be, in terms of absolute frequency, very rare indeed.[13] What matters is that they instantiate regular conjunctions of events (and hence make such events in principle predictable: see Introduction), which the accidental do not:

> **161** that there is no science of the accidental (*kata sumbebēkos*) is clear: for all science is either of that which is always or of that which is for the most part. For how else could one learn or impart it to another? Things must be determined as either occurring always or for the most part, e.g. that honey-water is for the most part beneficial to fever patients. But one will not be able to state when what occurs contrary to this happens, e.g. ‘at the new moon’—for then it will be the case on the day of the new moon either always or for the most part—but the accidental is the opposite of this. (*Metaphysics* 6. 2. 1027^{a}20–7)

‘Honey-water is for the most part beneficial to fever patients’ is true just in case, in most of the cases of fever in which honey-water is administered, patients recover as a result. ‘For the most part’, then, allows causal connections to be defeasible. The last sentence might be thought to hint at the possibility that ‘for the most part’ vaguenesses may be ultimately eliminated from science by further research that results in the formulation of more rigorous conditionals: propositions of the form ‘most *A*s are *B*s’ being replaced by ‘all *A*s, given conditions *C*, are *B*s’, where the content of *C* is clearly and independently specifiable. But that hint (if such it is) is never developed: and Aristotle rather suggests that there will always be some merely accidental cases of honey-water not being beneficial: and if they are merely accidental, then they are not themselves capable of formulation in universal or majority form (see *Posterior Analytics* 1. 30).

In fact, the reason why universality fails in the natural world is precisely that noted earlier: for Aristotle material facts get in the way of form realizing itself. There will be *individual* reasons why the form fails to come

end p.138

to fruition (this sapling was blasted by lightning; that one eaten by goats); but these cases are not, Aristotle thinks, in general susceptible of formulation in terms of some further generality (*Generation of Animals* 4. 4. 770^{b}9–17). This is why purely accidental conjunctions resist scientific explanation: there are no general relations of form under which they can be subsumed.

But for all that, they may still be necessitated in some way, and have causes:[14]

> **162** that there are principles and causes which are generable and destructible without there being for them any *process* of generation and destruction is obvious; for otherwise everything will be of necessity, since what is in the process of being generated or destroyed must have a cause which is not accidentally its cause. Will this occur or not? Yes, if such and such happens, but not if it does not. And this will occur if something else does. And so if time is constantly subtracted from a limited temporal extent, we will clearly arrive at the present. This man, then will die by violence if he goes out; and he will do this if he is thirsty; and he will be thirsty if something else happens. . . . For instance he will go out if thirsty, and he will be thirsty if he eats something spicy; and this is either so or not; so he will either of necessity die or not die. . . . Everything, then, that is to be will be of necessity; e.g. it is necessary that someone who is alive must some time die, since something has already occurred [sc. to necessitate it], namely the existence of contraries in the body. But whether he dies by disease or by violence is not yet determined, but depends on the happening of something else. Evidently then the process goes back to a particular starting-point, but this does not itself indicate anything further [in the past]. This, then, will be the starting-point of what happens by chance, and will have nothing else as the cause of its coming to be. (*Metaphysics* 6. 3. 1027^a29–1027^b14)

This is the closest Aristotle ever gets to discussing physical determinism (although *Nicomachean Ethics* 3. 5 bears comparison in the field of human action: §4*b* below), and scholars are divided over its interpretation and import.[15] But whatever else is true, Aristotle is clearly concerned here with the scope and origin of explanations. We know from the mere fact that you are human that you will die, decay and ultimate death being written into the formal structure of the human species. Furthermore, that structure explains the inevitability of your death: you die *because* you are human (or, more accurately, because you are an animal, and all animals are mortal: the proper syllogistic structure of explanation is examined at V.1*a*). But we cannot infer from that structure the time or manner of your death: that is not determined *by your form*.

Elsewhere (*Physics* 2. 4. 195^b30–2), Aristotle allows that chance and luck are indeed causes: but, as the first sentence of **162** insists, they are not causes for which there is any process of generation. A process in this sense is a continuous unfolding of form (in a relaxed sense of ‘form’): in the case of coincidences, there is no such unfolding, and hence no possibility of predicting them, at least from within the confines of each of the causal sequences that leads to them. To expand Aristotle's example, a predilection for hot food is the remote cause of your death, by making you thirsty, and thus encouraging you to go to the well and drink, where you happen to encounter ruffians who murder you. Those events are sequentially determined as part of a natural process, and to be explained (among other things) by your partiality for curry. But such partialities cannot be put in any general correlation with violent deaths, and hence cannot be part of any explanation of yours: when you happen to fall in with the ruffians, a new sequence of events is initiated, which does have a repeatable form (encounters with such people are liable to end in death): but the encounter itself has no antecedent process.

3. Teleological Explanation

(a) Teleology and Regularity

Equally, such accidents have no (genuine) final causes: you went to the well in order to drink, but not in order to die. Final causes are associated only with things which occur always or for the most part; and they are invoked precisely in order to explain those regularities. The basic intuition Aristotle is operating with here is that no description of the physical world that concentrates solely on material and efficient principles can suffice to account for the order and repeatability of natural physical processes. That is not to say that there is (metaphor apart) design or intentionality in nature. Rather it involves seeing particular physical processes (the maturation of a tree or infant, for instance) as being in a sense explanatorily basic. The mere material collocations allowed by the Atomists (VI.1*d*), and (on one view) Empedocles (I.4*c*), cannot account for this constant stability. Moreover, you misrepresent nature if you concentrate on the efficient and material aspects of its causal explanation at the expense of the others (cf. I.1*b*, **1–2**). At times, indeed, Aristotle's invocation of purpose in nature invites the conclusion that he does after all presuppose a Nature which consciously entertains its goals:

163 Nature, like an intelligent human being, always assigns each organ to something that is capable of using it . . . the most intelligent should be able to use

end p.140

efficiently the greatest number of tools: and the hand appears to be not so much a single tool as a tool of tools, as it were. Thus Nature has given the tool which is most widely useful, the hand, to the creature which is capable of acquiring the greatest number of skills. (*Parts of Animals* 4. 10. 687^a19–23)

Aristotle is here attacking Anaxagoras' view that humans are most intelligent because they have hands; rather (Aristotle thinks) men have hands because they are intelligent. And Aristotle elsewhere criticizes Anaxagoras (as Plato had done: III.1*a*, **122–3**) for not making sufficient use of his teleologically promising concept of Mind, but simply ‘using it as a stage device for introducing order when he is at a loss to explain why something is necessary’ (*Metaphysics* 1. 4. 985^a18–20).

Whatever we think of teleology of this sort, Aristotle is clearly justified in pointing to the range and multiplicity of the functionally adaptive features (as we would see them) of animals' bodies. Nature is only metaphorically a designer; and his constant insistence that nature does nothing in vain (e.g. *On the Heavens* 1. 4. 271^a33; *Parts of Animals* 2. 13. 658^a9; *Generation of Animals* 2. 5. 74^b13 etc.) is simply an affirmation of its end-oriented (as opposed to intentional) nature (as well as a heuristic to the isolation of function). Aristotle insists that formal and final causes are just as appropriate to biology as their efficient and material cousins. Consider antler-shedding:

> **164** deer are the only animals in which the horns are solid throughout, and are also the only animals to shed them, on the one hand for the sake of the advantage gained by the increased lightness, but on the other hand from necessity because of their weight. (*Parts of Animals* 3. 2. 663^{b}12–14)

Humans have cranial hair both as a result of material necessity (moist residue from the brain is excreted as hair) and for protection (ibid. 2. 14. 658^{b}2–6): nature 'invariably brings about the best arrangement of such as are possible' (2. 14. 658^{a}24). Teeth fall out necessarily (their roots are thin and easily work loose), but also to the animal's advantage, since they go blunt easily and hence require replacement (*Generation of Animals* 5. 8. 789^{a}8–15).

Elsewhere we discover that light passes through lantern-mesh both because of its fineness and 'in order that we may not stumble'; while if the Pythagoreans are right, thunder occurs not only of necessity when fire is quenched in the clouds but also to frighten the souls of the wicked in Tartarus (*Posterior Analytics* 2. 11. 94^{b}28–34). But these two are (at least allegedly) cases of consciously entertained purposes: the appeal to ends explains not how the lantern functions, but rather why it was constructed in the first place; and presumably it is Zeus' desire to frighten malefactors

end p.141

that moves him to extinguish nebular fire. Consciously intentional teleological explanation coheres with explanation in terms of material- efficient necessity without tension, since the explananda are in each case separate. Furthermore, what is causally efficacious is not the final goal as such but the antecedent desire of the agent to bring it about: there is thus no tendency to imagine that such an explanation involves causes exercising an influence backwards in time.[16]

(b) The Compatibility of Teleology and Mechanism

But matters are not so simple in the biological cases, shorn as they are supposed to be of intention. We need to examine just how far they can be assimilated to the intentional examples, and what function their appeal to ends is supposed to fulfil. These issues are sharpened by the objection that in the natural cases at least teleological and material-efficient explanations are actually incompatible.[17] The objection goes as follows. Suppose some explanandum *E*, and a set of conditions *C* which account for *E* mechanistically. It is then reasonable to suppose that the instantiation of *C* necessitates *E*. Now take some supposed final cause *F*: if it is going to be even part of the explanation of *E*, it should be at least a necessary condition of *E*. But the only way for that to be true compatibly with *C*'s necessitating *E* is if *F* is at least necessary for *C* as well. But how can final causes be necessary for the material instantiation of the mechanistic causes of things? Material-efficient and non-intentional teleological explanations of naturally occurring events are incompatible.

If the teleology is of the consciously directed kind, there is no problem: it is easy to see how the *conception* of the goal figures in the causal story which gets the means to that goal set in motion. But in the case of natural teleology, what can the appeal to ends accomplish over and above what is already done by the specification of materials involved? Teleological ‘explanation’ may be simply heuristic, as in the case of explaining the thermostat's function to a child in terms of its ‘desire’ to keep the room at a moderate temperature. But Aristotle's teleology is no mere heuristic device: he thinks that natural explanation without final causes is actually deficient. But what is missing from a purely mechanical account of antler- shedding, or one which explains the conception and generation of an organism solely in terms of the structure of semen and menses, and their interaction? The challenge is to interpret Aristotle's natural teleology

sufficiently strongly to allow some genuine explanatory power without courting the incoherence raised above.[18]

The issue is one of reduction. Why cannot Aristotelian natural purposes be seen as simply emergent on the physical structure of their material components? Reductionism may be ontological: we may assert, with Anaximenes, that heat really *is* rarefaction (I.2*i*). Or it may be methodological or theoretical: we may treat some theories as being reducible to other, more basic or general ones (thus Kepler's planetary laws reduce to Newtonian mechanics: they are seen as special cases thereof). Aristotle resists any ontological reduction of natural purposiveness: final causes are real components of the world.[19] Aristotle's teleology cannot, then, simply be a matter of supplying intuitively satisfactory answers to questions about the emergence of structure, since such ‘explanations’ are epistemic, and not in principle irreducible.[20]

So what stands in the way of our specifying properties of living organisms' material components which cause them to develop as they do? *Parts of Animals* 2 opens with an important statement of methodology:

165 the order of generation is the opposite of that of being: things prior in generation are posterior in nature, and the primary thing is last in generation (for a house does not come to be for the sake of the bricks and stones, but they for the sake of the house). . . . Everything generated makes its generation from something and to something, namely from one principle to another, from the primary motive principle which already has within it a certain nature towards a form, or some other end. For instance, man begets man, a plant a plant, each from their appropriate matter. Thus matter and generation are prior in time, but the substance and form are prior in formula. (*Parts of Animals* 2. 1. 646^{a}25–^{b}1)

At the most fundamental level of material construction there are the elements; they are followed on our ascent up the ladder of form by the uniform substances they compose, the non-uniform organs, and finally the complete creature (2. 1. 646^{b}2–10). Yet the fully actualized form is prior to its component parts (this is a general Aristotelian thesis: cf. *Politics* 1. 1–2. 1252^{a}–53^{a}38), which themselves exist for the sake of the form.

We can cash this out in at least two ways. We might suppose (weakly) that (*a*) the concept of parts presupposes that of the wholes they make up. But this linguistic fact (if it is one) has no obvious metaphysical implications. Aristotle's claim about the metaphysical priority of whole to part needs a stronger interpretation: (*b*) there could not have been parts (however they are to picked out) unless the wholes of which they are parts were to be generated. That metaphysical thesis recognizably assigns priority to the fully realized wholes (cf. *Metaphysics* 7. 11). But in ascribing to the whole a quasi-causal, genuinely explanatory role, it runs the risk of making the teleology and mechanism incompatible.

It has been suggested[21] that final causes are 'irreducible potentialities for form', irreducible just in that they cannot be attributed to the matter of which things are made up. Aristotle clearly requires something along these lines;[22] but it is quite unclear whether such a view can be made plausible. Given that suitable material-efficient conditions necessitate their outcomes, they will do so under any description: one which does not refer to those outcomes may mislead, but it will not be false. Only if it is the case that, in default of the final cause, these very same material-efficient conditions would not have generated the outcome has the final cause any genuine role to play. Yet that is precisely what is denied by the thesis of material sufficiency.

It is certainly true that simple piling the materials together in a heap will not, for Aristotle, generate complex organisms: but that is precisely where the organizing efficient cause comes in. The semen induces motions in the menstrual fluid, thereby inaugurating the embryological sequence (*Generation of Animals* 2. 1–2, 4. 1–3). Finality can appear in this story only by way of specifying the nature of particular efficient causes which are self- generating (man begets man, and plants, plants: **165**). Nature's processes will be goal-directed in the sense that we need to see how the various stages of growth and organogenesis contribute to the overall functioning of the fully actualized creature. But that does not appear to require that they be explained teleologically.

(c) Hypothetical Necessity

At this point the notion of hypothetical necessity[23] requires consideration:

166 is what occurs of necessity hypothetical or absolute? For people nowadays think that there is necessity in generation, just as if someone were to think that the wall came to be necessarily, because heavy things naturally sink, while light things rise to the top, so that stones descend to form the foundations, while the brick rises because of its lightness, with the timber at the very top since it is the lightest. But while it cannot come to be without these things, it does not do so because of them (except in the material sense), but rather for the sake of covering and preserving things. And equally in all other cases in which there is a for the sake of something: without things that have the requisite nature, but not because of them (except materially), but for the sake of something. For instance, why is the saw thus? In order to do this and for the sake of that. But this 'for the sake of' cannot occur unless it is made of iron. Therefore it is necessary that it be iron if it is to be a saw and perform its function; this necessity is, then, hypothetical, and not as an end is: for this necessity is in the material, while that for the sake of which is in the formula. (*Physics* 2. 9. 199^b33–200^a15; cf. *Parts of Animals* 1. 1. 639^b21–30, 642^a1–14, 33–6; *On Generation and Destruction* 2. 11. 337^b14–338^a4)

x is hypothetically necessary for *y* just in case if *y* is to come to be, *x* must be present or available. The materials are necessary for the building in that it cannot be built without them; but they do not in themselves necessitate its construction. Similarly, if something is to be a saw, it must be realized in a material capable of taking an edge: wool or wood will not do (*Metaphysics* 8. 4. 1044^a29); but merely producing a lump of iron will not make a saw. But appealing to the hypothetical necessity of certain materials is not meant to be an alternative to teleological explanation: rather it is part of it.[24] The required end determines that, in view of their physical properties, only a restricted set of materials can in fact be employed in realizing it; but they do not yet necessitate it. Atomism goes wrong in supposing that complex and regular outcomes can be accounted for solely on the basis of material necessities of the second type.

The distinction between hypothetical and absolute necessity points up sharply what Aristotle feels to be deficient about mere efficient-material explanation: no true description of the material *as materials* will be such as to entail, and hence explain, the development of complex structures out of them. But that is to assert anti-reductionism, rather than argue for it: and I can find no such argument, beyond an appeal to the intuition that matter on its own, and its merely physical interactions, could not produce the phenomenon of repeatable structure, and the view (succinctly expressed at *Metaphysics* 12. 6. 1071^b24–1072^a18) that material substance is, ultimately, passive and inert, always requiring something further to create and maintain movement within it.

On the other hand, there is something to be said for those intuitions, at least when contrasted with the explanatory resources of the Atomists. Mere haphazard atomic interaction does indeed seem inadequate to generate a stable, self-perpetuating world (see VI.1*d*, 2*a*, *d*; VII.3*a*), as Plato also held (III.4*d*). The Atomists' favoured basic properties of weight, resistivity, and solidity will need to be supplemented with something explanatorily richer, which can at least be conceptualized as the urge to

achieve form. We need to understand the concatenation of physical factors which are involved in the germination of seed, for example, as requiring description in terms of the eventual outcome (if nothing gets in the way: *Physics* 2. 8. 199^a11). No description of the materials alone will do, since they may be arranged in many different ways: yet each seed develops according to a particular plan (2. 8. 199^b10–14).

And this goes for other mechanistic theories as well. Aristotle sums up his opposition to Empedoclean mechanics as follows:

> **167** in general, anyone who says this destroys natural things and nature itself. For the natural things are those which move continuously in virtue of some principle within themselves towards a particular goal; (2. 8. $199^{b}14$–17)

in other words, to possess a principle of motion presupposes that that motion be goal-directed. At the end of the chapter, Aristotle emphasizes the internal nature of the form involved: nature does not deliberate about distinct ends (this is why we need not ascribe intelligence or planning to animals, much less to plants: 2. 8. $199^{a}21$–30). It is rather as if a doctor were to cure himself: for nature, agent, and patient are one (2. 8. $199^{b}30$–2). Final causes, then, are parts of reality in the sense that the drive for form that they represent is written directly into the structure of things. They are not ghostly, as-yet-unrealized objects exercising a mysterious *a fronte* causal power: rather they are the forward-looking elements of the incipient structure of organisms, a structure whose real existence allows Aristotle to reject the view that the Universe is controlled by the providential hand of a beneficent deity without thereby reducing it to what he at least sees as the absurd randomness of the pure mechanists (cf. VI.1*d*).

(d) The Scope of Teleological Explanation

We need still to determine the range of Aristotle's appeal to final causality. It is relatively easy to see how the function of parts can be explained in terms of their contribution to the overall working of the things of which they are parts; and those explanations are teleological. But it is more difficult and controversial to determine just what sorts of things are wholes. To be a whole, in the relevant sense, is not merely to be a collection or set; rather the whole must be, in some non-trivial sense, a unity; and one way of making sense of that notion of the unity of the whole is in terms of the parts as being contributory to its overall functioning. But again it is hard to achieve any such characterization in clear, non-question-begging terms, since the extent to which one is inclined to see the parts of the Universe as contributing to it as a whole in this sense

will depend to a large extent on how far one is already prepared to see it *as* a whole. Aristotle does indeed view the Universe as a unity (*On the Heavens* 1. 1–4), and indeed partially derives some at least of the properties of the elements from its overall structure: the innate capacities of fire to rise and earth to fall rely on a notion of the structure of the Universe as a whole in order to give content to the concepts of natural place which they employ (ibid. 1. 8; §3*b* below).[25] Indeed what seems to a modern reader to be the excessive aprioricity of Aristotle's cosmology is driven by this fundamental assumption, and the related thesis that in the metaphysical order of things the completed whole is prior to its component stages, phases, and parts (cf. *Metaphysics* 12. 10).

This suggests that Aristotle may view individual things within the cosmos not merely as wholes in their own right, towards whose activity the proper order and functioning of the parts is contributory, but also as parts of a larger whole, a biosphere in which the lesser contributes to the welfare of the greater. In the notorious chapters of the *Politics* which describe the condition of natural slavery, Aristotle implies that the lower animals (in particular domestic animals) exist for the sake of human flourishing (*Politics* 1. 5. 1254^{b}10–12, 8. 1256^{b}10–22): if nature does nothing in vain, we must conclude that plants exist for the sake of animals, and animals for the sake of men. None the less, the point is not so much that Nature ordained their use by man; rather they are actually *better off* under human control, since men (at least those who are not by nature slaves) are best fitted to rule. They can only achieve their own goals in life as part of a community in which they benefit human beings. Even so, it requires the eye of faith to discern such providence in action (particularly if you happen to be a pig).

The notion that nature is structured so that its higher inhabitants benefit from the lower is not necessarily anthropocentric, although such views tend naturally in that direction. The claim might be purely descriptive (indeed, it might form part of a definition of 'higher' in the appropriate sense). Moreover, what has been established already will naturally tend to discredit any consciously anthropocentric teleology in Aristotle: the forces of nature are not blind, in the sense that they are not merely the products of Democritean chance; but they are not providential either.

On the other hand, a passage of the *Physics* 2. 8 has recently been re-evaluated in an anthropocentric fashion.[26] The purpose of the chapter

end p.147

is to say 'why nature is one of the causes for the sake of which' (2. 8. 198^{b}10–11):

> **168**(i) there is a difficulty here: why cannot nature act not for the sake of anything or because it is better, but rather in the way Zeus rains, not in order to make the corn grow, but of necessity? (ii) For what rises must become chilled, and what is chilled becomes water, and then falls, while the growth of the corn supervenes upon this event. (iii) Similarly, if someone's corn is spoiled on the threshing-floor, it does not rain for the sake of this, to destroy the corn, but this just happens to occur. (iv) So what prevents the parts in nature from being like this, for instance in the case of teeth it is necessary that the front ones are made sharp and fit for cutting, while the molars are broad and useful for grinding up the food, it not being for the sake of this that they became like this, but simply coincidentally? (v) And similarly with the other parts that seem to exist for the sake of something: wherever everything happened to occur as though it were for the sake of something, these were preserved, having been suitably formed by chance; while those which were not thus perished and still perish, as Empedocles says of the man-faced oxen. . . . (vi) But things cannot be so. For these things and everything that occurs by nature do so either invariably or for the most part, while those things that happen by luck or chance do not. (vii) For it does not seem to be by chance or coincidence that it rains frequently in winter, but it does in August (neither does heat so appear in August—but it does in winter). (viii) So if it seems to be either coincidental or for the sake of something, and if these things cannot be either coincidental or chance, they will be for the sake of something. (ix) But these things [i.e rain in winter, heat in summer] are natural, as even the people who say these things agree. (x) Therefore there is purpose in things which come to be and exist by nature. (2. 8. 198^{b}16–199^{a}8)

The crucial interpretative crux surrounds the question of whether or not Aristotle endorses the claim that rain is not for the sake of anything (i–ii). Clearly he rejects the purely mechanical account of animals' teeth (iv–v): they exist both for the sake of eating and as weapons (*Parts of Animals* 3. 1. 661^{b}7–662^{a}16): does he then intend to commend a purely mechanical account of rainfall? Certainly Aristotle would endorse the mechanical account he offers as far as it goes (*Meteorology* 1. 9; cf. *On Sleep* 3. 457^{b}31–3); but as we have seen, Aristotle does not claim that material-efficient explanations are false, only that they are inadequate. Moreover, (vii–x) offer an argument for the teleological nature of winter rain: it happens regularly, no regular occurrence is a matter of chance (cf. **160–2**), but anything which is not a matter of chance must be for the sake of something (it is of course a further question whether winter rain is for the sake of crop growth).[27]
end p.148

The crucial step in the argument is (viii). Unless we accept that the disjunction ‘either chance or for the sake of something’ is exhaustive, we need not be forced to the conclusion that winter rain is purposive (and hence by (ix) that some natural things exhibit purpose). We have already seen (§2*b* above) that Aristotle assimilates the explicable to what happens always or for the most part; and this is opposed to what happens by chance. Chance occurrences are not (then) fully explicable: and they lack final causes. But this does not in itself entail that anything that is explicable must have a final cause: might there not be absolutely regular yet purely mechanical operations, such as, for instance, the motions of the heavenly bodies? Yet Aristotle does think that they exhibit final causality: they move as they do in order best to imitate the nature of the Prime Mover (*Metaphysics* 12. 7–10: V.2*g*). Whatever we are to make of that dark doctrine, Aristotle seems to have supposed that everything strives, in so far as it can, to emulate the First Cause (thus mortal things seek immortality of a sort in reproduction: *Generation of Animals* 2. 1. 731^{b}18–34; *On the Soul* 2. 4. 415^{b}1–7; this idea is ultimately owed to Plato: *Symposium* 207c–8b). Thus everything has (in some sense at least) a purpose (see §4*a* below).

(e) Explanation and the Good

Purposes, however, can be expressed in different ways. *Parts of Animals* is a treatise in philosophical biology, devoted to the analysis of the relations that hold between animals' parts and their overall structure. More particularly, in line with Aristotle's unconcerned acceptance of the compatibility of teleological and mechanical explanation (§ *b* above), it seeks to provide a complete account of animals' structures in terms of both physical necessity and functional appropriateness. We saw that antler-shedding could be given both material and final causes (*Parts of Animals* 3. 2. 663^{b}12–14: § *a*, **164** above); that account comes in the middle of a general discussion of the various ways in which nature has provided for animal's defence, assigning sharp claws to some, sharp teeth to others, and so on (3. 2. 662^{b}23–664^{a}12); all of these have functional explanations, and while nature goes about protecting different animals differently, it never supplies more than one

‘sufficient defence’ to any animal (3. 2. 663^{a}17–18; cf. 662^{a}32–4): nature is thus economical. But equally, he argues that nature's arrangements are providential: the best place for horns is (contrary to what Aesop said)
end p.149

on the head (3. 2. 663^{a}34–^{b}12), where they offer maximum protection for minimum encumbrance.

Even so, sometimes horns are without purpose (3. 2. 663^{a}8–11; they can even be a disadvantage: 2. 16. 659^{a}19). The reason for this is that they also have a material cause (3. 2. 663^{b}23–664^{a}12):

169 as there is a residual excess of this type of matter [i.e. earth] in the larger animals, nature turns to protection and advantage, allotting that part which of necessity flows upwards to teeth and tusks in some animals, horns in others. (3. 2. 663^{b}32–6)

Aristotle makes much use of the concept of residues (*perittōmata*) in his explanation of organogenesis. Residues are the necessary by-products of essential metabolic processes (2. 2. 647^{b}27–9) involving the uniform substances that supply nutrition (as distinct from those which form the material basis of the non-uniform organs: 2. 2. 647^{b}21–5). Such residues will be put to good use by nature (as in the case of horns), if circumstances allow; but on occasion there will be residues which cannot be turned to any purpose. These will remain residues pure and simple, for which there is no direct teleological explanation. Yet they are the necessary results of processes which are themselves goal-oriented; and so even here the universal grip of teleological explanation is not fully relaxed. Thus the spleen is a necessary residual outcome of the functioning of the stomach and bladder (3. 7. 670^{a}31–^{b}17), and does not invariably fulfil any function (although it may assist in metabolism). By contrast, the heart and the liver subserve some direct function (3. 7. 670^{a}23–30).

Moreover, some organs are not the result of necessity at all, but are there solely to improve some function. Such an organ is (Aristotle thinks) the kidney (3. 7. 670^{b}23–8): it exists (where it does) to assist the function of the bladder. The necessity in this case is hypothetical, not material (§3*c* above): the functioning of the animal does not absolutely require a kidney; but the presence of one will, in some cases, help.

We are now in a better position to turn to the issue raised at the beginning of this section. Final causes ‘can be the good for some thing, or of something’ (*Metaphysics* 12. 7. 1072^{b}1–2). This obscure remark is lent some clarification elsewhere (*Physics* 2. 2. 194^{a}33–6). *A* may supply a final- cause explanation for *B* in the sense (i) that *B* acts for the benefit of *A*. This is the basic structure of the relation in the biological cases. Bones are for the sake of flesh (i.e. to hold it in place: *Parts of Animals* 2. 8. 653^{b}33–5, 2. 9. 654^{b}27–655^{a}1). Ultimately the whole structure of an animal's body will be largely explicable in terms of the manner in which its component parts are conducive to the functioning of the animal's soul, its proper form of life (see V.3*b*). However, A may also be final for B in the sense (ii) that B aims at A, without in any way being hypothetically necessary for A's existence.
end p.150

This is the relation that holds between lower and higher denizens of the scale of nature (§ *d* above; cf. §4*a*, **174** below; V.2*g*, **203**).[28]

It remains to add a few remarks on one of the most distinctive, not say problematic, features of Aristotle's teleology: his appeal to the notion of the honourable in order to justify certain arrangements teleologically (such appeals surface in cosmology too: V.2*g*, **200**). In *Progression of Animals* 4–5, Aristotle discusses these arrangements: all living things have 'upper' and 'lower' parts, where these are determined functionally, and not simply by relative position. The 'upper' parts are those 'from which come the distribution of nourishment and growth' (*Progression of Animals* 4. 705^{a}29–32), and hence (in this sense) plants are upside down (4. 705^{b}6–8). Additionally, animals have a front and a back, which is also determined functionally, by the location of the sense-organs (4. 705^{b}8–14). Finally, animals capable of locomotion are also distinguished by left and right, again defined functionally (the right side is where the origin of motion derives, the left that which naturally follows: 4. 705^{b}14–21). Of course, many creatures do not exhibit their 'true right' side on the actual right- hand side (crabs are particularly deficient in this regard: 14. 712^{b}13–22, 16. 713^{b}12–16); however:

170 human beings have their left parts distinguished more than any other animal on account of their being the most in accordance with nature of any animal, since by nature the right is better than and separated from the left; and for this reason the right sides are most particularly right in humans. . . . And the other sources (*archai*), the upper and the forward, are also particularly differentiated in accordance with nature. (4. 706^{a}17–26)

Moreover,

171 bipeds have it [i.e. the 'upper' part] at the top, since they are upright, human beings in particular, since they are the bipeds most particularly in accordance with nature. And it is reasonable that the sources should be from these parts, since the source is honourable, and the upper is more honourable than the lower, the forward than the backward, and the right than the left. It is equally correct to assert the converse concerning these things, that because the sources are in them these are more honourable than the opposing parts. (5. 706^{b}9–16)

That last sentence perhaps suggests that what makes the preferred locations more honourable is simply that they do contain the 'sources';[29] yet Aristotle must also think that there is something *naturally* honourable about these locations, independently of their containing the superior functions, since otherwise there is no way to give content to the suggestion that humans are the most naturally constituted of all species.

Indeed, this principle is put to work in a complex piece of reasoning designed to vindicate nature's organization of oesophagus and windpipe (*Parts of Animals* 3. 3). At first sight, the arrangement seems botched, since, because the former is located behind the latter, food must pass across the windpipe before entering the oesophagus, a fact which makes choking a possibility (3. 3. 664^{b}2–7), and which necessitates (hypothetically) the epiglottis (3. 3. 664^{b}19–36), at least in moistly constituted animals who have an oesophagus. The oesophagus is itself necessitated by the lungs,[30] which have to occupy the place they do in order to fulfil their function of cooling the heart, whose place is itself determined by the demand that it, as the ultimate source, should be located in the most honourable position:

172 let these things be said, then, concerning the reason why some animals have and some animals lack the epiglottis, and that nature has patched things up because of the poor position of the windpipe by contriving what is called the epiglottis. The larynx lies in front of the oesophagus of necessity, since the heart lies in front and in the middle, in which we say is the source of all movement and perception for the animal (for both perception and movement are towards that which is called the front, since front and back are distinguished by this principle (*logos*)). The lung lies where the heart is and around it, while breathing occurs because of it and because of the source which inheres in the heart. Breathing occurs for animals through the windpipe. Consequently, since it is necessary that the heart lie first of all in front, both larynx and windpipe must lie in front of the oesophagus: for they lead to the lung and heart, while it goes to the stomach. In general, wherever nothing more important prevents it, the better and more honourable is always in the upper parts of the upper and lower, the front of the front and back, and the right of the right and left. (3. 3. 665^{a}7–25)

The nobility of the heart's position is reaffirmed in the next chapter (3. 4. 665^{b}18–21); and everything else in the arrangement is subordinated to its claim on the superior position (cf. 2. 14. 658^{a}22–4). Equally, the diaphragm serves to divide the lower parts (the stomach and digestive system) from the upper (the heart, as seat of sensation and cognition), with the former existing by (hypothetical) necessity for the sake of the latter (3. 10. 672^{b}9–24).

The argument of **172** is at the very least ingenious: and it does serve, in a sense, to explain the existence of what might otherwise seem a somewhat Heath Robinson contraption, the epiglottis. Moreover, there is nothing intrinsically problematic about considering the various functions of the body to be hierarchically organized, with the superior functions being

those to which the others are subservient. And, in the manner of the last sentence of **170**, one might choose to consider the superior functions (and by association their locations) as being more noble. Yet it seems as though for Aristotle the location itself must possess nobility independently of what is located in it, otherwise Aristotle has no explanation for the arrangement of windpipe and oesophagus (since if this were not the case the heart could have been located behind, thus making that location more noble). Aristotle might have attempted an argument based upon the fact that the organs of perception must, if they are to be efficient, be located facing the natural direction of travel; but he does not, in this context, try to do so. Moreover, it is not easy to reconcile such an active role for the principle of nobility from within the framework of a non-intentional, undirected teleology: *how* can nature aim at what is more honourable? Such worries are well-motivated; and they were to worry Theophrastus. But that is a story for the next chapter (V.3*a*).

4. The Explanation of Action

(a) The Soul

Living things, particularly animals, are paradigms of teleological organization; and those organizations are, for Aristotle, their souls. Aristotle rejects Plato's account of a tripartite soul (*Republic* 4. 435a–442e) in favour of an account which stresses the existence of different psychic faculties: metabolic and generative (associated with the ‘vegetative soul’), perceptive and locomotive, and rational. Soul is just, for Aristotle as for the Greeks in general (cf. I.2*c*, 3*c*), what makes living things alive:

173 we say, then . . . that the ensouled is distinguished from the unsouled in that it is alive. But ‘alive’ has many senses, and we call something alive if even one of these things inheres in it: mind, perception, motion and rest in regard to place, and furthermore the motion in regard to nutrition, and decay and growth. For this reason all plants seem to be alive, since they clearly possess in themselves a power and a principle of such a kind through which they enjoy growth and decay . . . and they are nourished and continue to live as long as they enjoy nutrition. But this can be separated from the others, while the others cannot exist without this in mortal things. (*On the Soul* 2. 2. $413^{a}20$–33)

In fact, these psychic functions form a hierarchy, the higher of which presuppose the lower: thus everything which is alive has the nutritive, metabolic faculty; the lowest animals have sensation only to the extent of touch (2. 2. $413^{b}2$–22; cf. 3. 12), but they too have the nutritive soul; higher animals are endowed with the other senses and the capacity for locomotion
end p.153

as well; while man is capable of rational thought (deliberation and abstract reasoning) as well as all the other things (2. 3. $414^{a}29$–$^{b}19$, $^{b}28$–$415^{a}13$).[31] The distribution of these faculties is to be explained teleologically. There is no point in anything being able to locomote unless it is capable of desire (and aversion); and it will only be so capable if it can feel pleasure and pain, i.e. if it has the capacity for sensation (3. 12. $434^{a}26$–$^{b}9$; cf. 3. 11. $434^{a}1$–3): nature does nothing in vain, and assigns faculties only to those creatures capable of utilizing them (3. 9. $432^{b}13$–19; cf. §3*a*, **163** above).

But even the lowest forms of life aspire to mimic the perfection of the Prime Mover to the best of their limited abilities (cf. §2*d* above; V.2*g*, **203**):

174 the nutritive soul belongs to the others, and is the primary and most common power of the soul, in virtue of which life belongs to all of them. Its functions are generation and the use of nourishment. For this is the most natural of all animal functions, at least for those which are complete and not deformed . . . namely to produce another like itself, animal to animal, plant to plant, so that they might participate in the eternal and the divine in so far as they can. For every creature desires this, and does what it does naturally for the sake of this. ('For the sake of' has two meanings [cf. §3*e* above]: that for the purpose of which, and to the benefit of which.) So since they cannot share in the eternal and divine by continuity, on account of the impossibility of mortal things remaining one and the same in number, they share in it in the way they can. (2. 4. 415^{a}23–^{b}6)

Every creature strives for immortality by way of reproducing itself (cf. *Generation of Animals* 2. 1. 731^{b}20–732^{a}9).

But what is the soul? Aristotle's account is analytical, rather than psychophysical: he is concerned to isolate the features of animal activity which we ascribe to various psychological and physiological capacities. The soul is the form of the living body (*On the Soul* 2. 1); and it is its first principle and cause, (*a*) efficiently, (*b*) finally, and (*c*) formally (2. 4. 415^{a}9–28). It is (*a*) the source of autonomous movement for the animal; it is (*b*) that for the sake of which the body is constructed the way it is; and (*c*) in a strong sense it *is* the living body: it is what distinguishes an animal from a mere heap of blood, bones, and flesh. Thus, while fire is, for Aristotle, a *sunaition* of nutrition (since fire is self-nourishing, and living things are warm), the proper cause is the (nutritive) soul itself (2. 4. 416^{a}10–18). The account of (*a*) is spelled out in more detail in *On the Soul* 3. 9–11. The two causes of movement, at least in the higher animals, are mind and desire (3. 10. 433^{a}10–13); but of these, desire is most properly motive, since there is no motion without it (433^{a}14–31). Desire pursues

what seems to it to be good, something which it represents to itself by way of imagination:[32]

175 what causes movement is one in form, namely the desiderative *qua* desiderative, and first of all the object of desire (for this causes movement without being moved, by being conceived or imagined), but the things which cause movement are numerically many. There are then three things: that which causes motion, that with which it moves, and that which is moved. The motion-causer is double, one of which is unmoved, the other both motive and moved. That which is unmoved is the practical good, that which is both motive and moved the desiderative faculty . . . and the thing moved is the animal. . . . In general, then, as has been said, an animal is mobile in so far as it is desiderative; and it is not desiderative without imagination. All imagination is either calculative or perceptual, of which latter the other animals share in too. (3. 10. 433^{b}10–31)

The 'calculative imagination' is that which allows human beings to deliberate, by producing mental constructions of possible future courses of events between which they can make choices, choices which both reflect and determine their characters, and in virtue of which weakness of the will, the preference for immediate over long-term gratification, becomes theoretically possible (3. 10. 433^b15–10, 3. 11. 434^a6–14). Animals, on the other hand, are limited to the form of imagination which allows them to evaluate, as worthy of pursuit or avoidance, only the immediate objects of perception. But all are inspired to move by the object of their desires, which excites motion in them by way of their desiderative faculty, but which does not itself move (in line with Aristotle's strictures on the impossibility of self-motion: see V.2*g*).

Moreover, the hierarchical structure of the soul (and the ultimate relationship of everything with the Prime Mover: §2*d* above; V.2*g*) enable Aristotle to define what he at least takes to be the proper function of human beings, a notion which is fundamental to the account of the good life he develops in the *Nicomachean Ethics* and the *Politics*. What defines an animal species is what it is able to do either uniquely, or better than anything else (*Nicomachean Ethics* 1. 7. 1097^b23–1098^a17), or perhaps rather the highest type of activity of which it is capable (10. 7). Thus it turns out that the function of human beings is to deliberate (in such a way as to enable them to organize in complex political and social units), and engage in abstract thought. These capacities are what set humans above the other animals; and in the latter at least they approach the perfect condition of God, the Prime Mover. Happiness is 'activity in accordance with excellence', and the highest human excellence is the ability to think abstractly

(10. 7. 1177^a11–23; cf. in general 10. 6–9). Human beings cannot think such thoughts continuously and without respite; but in so far as they are capable of it, in doing so they most fully realize their natures, and come to touch the divine.

(b) Freedom and Responsibility

Animals move autonomously, in virtue of their own internal structures. But those movements are conditioned, and perhaps ultimately caused by states of affairs beyond their immediate control. The structure of human action is more complex than that of other animals, involving as it does higher-order forms of imagination (**173**; cf. *On the Soul* 3. 7) and deliberation (*Nicomachean Ethics* 3. 3). But does that complexity make us in any strong sense free of causal constraint? And if it does not, what happens to our ordinary conceptions of autonomy and responsibility? The challenge is issued by Gorgias'*Helen* (II.2*c*): if our actions are indeed determined by factors beyond our control, how can we be responsible for what we do? Gorgias is the founding father of hard determinism. Aristotle opens his discussion of voluntariness and responsibility as follows:

176since ethical virtue is concerned with emotions and actions, and those which are voluntary are praised or blamed, while those which are involuntary receive pardon and sometimes pity as well, students of ethical virtue must presumably determine the limits of the voluntary and the involuntary. . . . Actions are regarded as involuntary when performed under compulsion or through ignorance. (*Nicomachean Ethics* 3. 1. 1109^b30–1110^a1)

Aristotle saw the consequences of attempting to mitigate responsibility by blaming factors beyond the agent's control:

177 if it were argued that pleasurable and admirable things have a compulsive effect (because they bring external pressure to bear on us), it would make all acts compulsory, since every act of every agent is done for the sake of such objects. . . . It is absurd for the agent to lay blame on the external factors and not upon himself for falling easy prey to them, and to attribute his fine acts to himself, but his disgraceful ones to the attractions of pleasure. It seems reasonable, then, that an act is compulsory only when its originating cause is external, and receives no contribution from the person under compulsion. (3. 1. 1110^{b}9–17)

We are responsible for what we do whenever the action originates within us; moreover, even if forced to do something we are (partly at any rate) responsible for it as long as we acquiesce in it. We are praised or blamed for the state of our moral characters; and we are morally accountable for actions undertaken on the basis of choice and deliberation. When we choose to do something, that choice is up to us, since it is the outcome of

our desires, beliefs, and deliberations: moreover, virtue itself is up to us (3. 5. 1114^{a}33–^{b}25). Helen may indeed have been persuaded by clever arguments, or swept along by passion: but it was *she* who was so persuaded and affected, and those facts tell us something about her state of character. Even if Gorgias' disjunction is exhaustive, Aristotle can still resist concluding that Helen's action was not her fault.

We obviously do hold people to account for what they do, and for that to be reasonable some things must be up to us. Aristotle's basic contention is that we are responsible for what we do as a result of our characters. You are commended for your courage in staying at your post under fire, while I am castigated for my cowardice in running away; and that fact reflects upon our different moral natures:

178 everyone aims at what seems to him to be good, but over this appearance he has no control. How the end appears to each individual depends on the nature of his character, whatever this may be. So if the individual is in a way responsible for his state of character he will also be in a way responsible for his view of what is good; but if he is not responsible for the former, then no wrongdoer is responsible for doing wrong. (3. 5. 1114^{a}31–^{b}4)

If virtue is simply a matter of natural endowment, or correct upbringing, it is hard to see how the acquisition of it can be in the individual's power, at least if responsibility is transitive. If we can reasonably hold individual *I*'s character *C* responsible for his action *A*, then, if responsibility for *C* itself rests upon some set of factors *F* outside *I*'s control, then *F*, rather than *I* himself, appears to be truly responsible for *A*.

That apparently plausible argument reopens the case for Helen's de- fence; and, as Aristotle realized, it is only if we are responsible for the development and persistence of our states of character themselves that we may legitimately be held to book for how they make us behave. Aristotle thought that our characters were originally at least sufficiently malleable for us to be able to influence the direction of their evolution. The performance of good actions makes us good, by creating in us a disposition to perform them (2. 4. 1104^{a}11–1105^{b}18). When the disposition sets hard, we cannot resist its injunctions; but it has not always been so concrete. And even if I am now so far gone in vice that I literally cannot prevail against it, there was a time when I was relatively immune to its siren song, and could choose to pursue it or to avoid it.

But when I 'freely' choose vice over virtue, why assume that my choice is somehow free of the constraints which later condition my vicious lifestyle? The currently fashionable distinction between first- and second- order desires does no work here (was not my second-order choice to develop a vicious disposition *itself* the result of existing traits of character, or external pressures?). What is more, Aristotle lays great store by the
end p.157

importance of a proper upbringing in developing a good character: but that, surely, is by definition outside the control of the individual who is being properly brought up. Moreover, to what extent (if at all) must we be free of the causal constraints exercised by the Universe? We deliberate only about things within our power, and for that to be the case, we must be able to affect the course of events. In a famous, and famously difficult, chapter of *On Interpretation*, Aristotle argues that the future must be contingent, otherwise deliberation itself would be pointless.[33] But it is not clear how Aristotle conceptualized this contingency,[34] and indeed whether it need amount to a rejection of determinism as such (cf. §2*b*, **162** above). But at the very least, Aristotle apparently thinks that if the world is now such as to make a certain prediction true, then deliberation about the content of that prediction (ways to prevent it, and so on) makes no sense (*On Interpretation* 9. 18^b26–19^a6; Epicurus apparently thought so too: VI.3*a*). There must exist a genuine contingency, which is defined as it being the case that things are no more likely to turn out one way rather than the other (9. 18^b6–9).

These considerations open up large questions which are better put off until later (VI.3*a*–*b*; VII.2*a*–*e*; X.4*c*; XI.5*c*). Suffice it to say here that all that is apparently required to lend deliberation point is for the future not to be fatalistically determined, i.e. structured so that a particular outcome is inevitable no matter what you do, or whether or not you deliberate. Even in a thoroughly deterministic universe deliberation may be causally efficacious: the deliberations themselves may partially cause the none-theless-determined outcomes. But even so, the question remains open whether, and if so how and to what extent, human deliberators can be responsible for their actions in such a world, a question that was to loom large in the later history of Greek philosophy.[35]

5. Conclusions

Aristotle was sympathetic to Plato's view that mere mindless mechanism could not generate a world (III.1*a*); but he rejected both Plato's metaphysics of hypostasized perfect Forms and his divinely directed teleology (§1*a*), in favour of a notion of immanent form in the essences of natural things, an
end p.158

internal, goal-directed, principle which drives living things towards maturity by organizing their material elements into the appropriate structure. Matter corresponds to potentiality, and realized form to actuality; things come to be not *ex nihilo* (in contravention of the Parmenidean strictures: I.3*d*), but out of pre-existing matter, under the influence of directing, internal form (§1*b*). Aristotle famously distinguishes four 'causes' (or causal factors in explanation), the matter, the form, the end, and the agent (§2*a*), and discriminates sharply between regular, natural events and those brought about by chance, which resist full or proper explanation (§2*b*).

Aristotle's teleology is invoked in biology to explain the links between structure and function (§3*a*), making goal-directedness an irreducible component in any satisfying explanation of the structural regularity of the natural world, although not at the expense of material explanation (§3*b*): the ends of nature can only be realized in particular types of suitable matter, which are thereby 'hypothetically necessary' for those realizations (§3*c*). Teleological considerations are closely linked, for Aristotle, with the idea that the parts are for the sake of the whole, and contribute to its overall good; and Aristotle suggests that, in some sense, the whole structure of the cosmos is teleologically organized (§§3*d*–*e*; cf. V.2*g*), a structure that finds expression in Aristotle's functional concept of the different degrees to which soul is exhibited in living things (§4*a*).

Finally, Aristotle confronts an issue which is to be of increasing importance in subsequent philosophy: how, and to what extent, can we justify our intuitions that human action is, from the point of view of moral ascription, free, given that we do what we do as a result of dispositions induced in us by external factors (§4*b*)? Aristotle's sketch of an answer picks up an earlier thread: human beings, as living things, have internal natures: action on the basis of those natures can be ascribed to us, particularly if we have developed the dispositions we have on the basis of earlier choices.

V Aristotle: Explanation and the World

R. J. Hankinson

1. Demonstration and Explanation

(a) The Structure of Science

So far, we have examined Aristotle's attitude to the scope and applicability of the various explanatory categories. Now we turn to his account of the logical structure of explanation in *Posterior Analytics*. Sciences, for Aristotle, are organized bodies of knowledge; and, at least if they are properly developed, their organization will reflect that of the portion of reality they describe. Certain truths about the world are simply basic, requiring no further explanation (*Metaphysics* 4. 4. 1006^{a}6–8; cf. *Posterior Analytics* 1. 3; *On the Heavens* 2. 5. 287^{b}29–32). Each science proceeds by treating them as axioms from which to derive the dependent theorems that form the remainder of the science. The deductive structure of the science will thus mirror the hierarchical structure of the world. Axioms must be

> **179** true, primary, immediate, more intelligible than, prior to and explanatory of the conclusion. . . . The premisses must be (*a*) true, since what is false cannot be known. . . . They must be (*b*) primary and (*c*) indemonstrable, since other wise they will not be known without demonstration; for to know that of which there is a demonstration (other than incidentally) is to have that demonstration. They must be (*d*) explanatory, (*e*) more intelligible, and (*f*) prior: explanatory since we only have scientific knowledge when we know the explanation; prior since they are explanatory; and already known not only in the sense that they are understood, but in that they are known to be true. (*Posterior Analytics* 1. 2. 71^{b}19–33)

(*a*) is uncontroversial. (*b*) and (*c*) come to the same thing: a proposition ('all *A*s are *B*s', or as Aristotle puts it, '*B* belongs to all *A*': symbolically, '*B*a *A*') is primary just in case there are no further propositions of which it is a deductive consequence, and which serve to explain it. As Aristotle would put it, there is no middle term '*C*' such that all *A*s are *C*s and all *C*s are *B*s (where *A*'s being *C* explains its being *B*): but that is precisely what it is for a proposition to be immediate. Aristotelian sciences are self-contained: they do not share axioms with one another (ibid. 1. 7. 13); thus it is not enough simply for an axiom to satisfy (*b*); in a given demonstration, it must also satisfy (*f*), i.e. be prior to the conclusion in question. (*d*)

end p.160

and (*f*) are also linked: some property *F* is prior to *G* just in case something's being *G* is explained by its being *F*. Finally, by 'more intelligible', Aristotle means more basic in the order of being: he distinguishes propositions that are more intelligible to us (i.e. more immediately and directly available) from those which are more intelligible in themselves, the bedrock propositions of science which concern him here.[1]

Aristotle's logic of explanation depends upon the syllogistic worked out in the *Prior Analytics*.[2] For the purposes of the *Prior Analytics* analysis, syllogisms are two-premiss arguments in which the two terms involved in the conclusion are mediated in the premisses by a middle term. For demonstrative science, the most important such structure is the paradigmatic first figure mood of Barbara: all *B*s are *A*s; all *C*s are *B*s; hence all *C*s are *A*s (which symbolizes as '*A*a *B, B*a *C* ⊢ *A*a *C*'); although sciences will tolerate, sometimes perhaps even require, syllogisms in other moods, even in other figures.[3]

A demonstration, for Aristotle, is a proof in Barbara in which the premisses (and hence the conclusion) are necessary truths; but not every such argument is, in the strictest sense, a demonstration, since there can be cases of valid arguments in which the premisses are necessary truths but which are not yet demonstrations. Consider

[A] (i) all mammals are mortal;

(ii) all men are mammals;

hence

(iii) all men are mortal.

[A] has the right basic form, and it deduces a necessary conclusion (at least let us suppose so) from necessary premisses. But it does not satisfy (*c*),

end p.161

since [Ai], the major premiss (i.e. the premiss which contains the predicate of the conclusion) is not, in the requisite sense, immediate. For a proposition of the form '*A*a *B*' to be immediate there must be no (genuine and non-trivial) *C* such that *A*a *C* and *C*a *B* are both true. That is, in the order of things, the connection between *A* and *B* must be tight. In the case of genuine demonstrations, some of the axioms will be definitions: and definitions convert (although not of course *qua* definition—if *B* defines *A, A* cannot define *B*): that is, if $A =_{df.} B$, then *A*a *B* and *B*a *A* both hold—any definiens must be coextensive with its definiendum. Moreover, some premisses will be partial definitions, that is one of their terms will figure in the definition of the other: Aristotle calls such predications *per se* (*kath'hauto*) predications (*Posterior Analytics* 1. 4. $73^{a}34–^{b}5$).**[4]** But at any rate, for a premiss *A*a *B* to be immediate, there must be no higher-level generalization which in turn explains why *A*a *B* holds; hence the all-important condition (*d*).

A demonstration explains why a particular predicate holds of its subject. That is to be done (at least provisionally) by selecting the widest genus for which it is true that the predicate holds of that genus, and of which the subject is a member. Thus an explanatory syllogism for the conclusion of
[A] might run as follows:

[B] (i) all animals are mortal;
(ii) all men are animals;
hence
(iii) all men are mortal.

According to [B], men are mortal in virtue of the fact that they are animals; their being animals, then, explains (in the sense of pointing to the feature responsible for) their mortality.

At least, that is what Aristotle sometimes suggests; but he wavers on the issue of which middle term in a sequence is genuinely explanatory. Consider a chain of demonstrative syllogisms: *A*a *B, B*a *C* ⊢*A*a *C*; *A*a *C, C*a *D* ⊢*A*a *D*; *A*a *D, D*a *E* ⊢*A*a *E*. There are three middle terms, *B, C*, and *D*, involved in proving that *A* holds universally of *E*, and three distinct syllogisms can be constructed to link *A* and *E*, namely: *A*a *B, B*a *E* ⊢*A*a *E*; *A*a *C, C*a *E* ⊢*A*a *E*; *A*a *D, D*a *E* ⊢*A*a *E*. Which of them is genuinely explanatory?
end p.162

That is, which of *B, C*, and *D* explain why *A* holds universally of *E*? In a sense, that question is absurd: surely they all do, since they are all links in the chain from *A* to *E*. But at *Posterior Analytics* 2. 18, Aristotle explicitly opts for *D*, the middle term 'closest to the subject'.

Yet in the previous chapter, Aristotle at least hints that things should go the other way (in line with [B]).[5] He is discussing the relations that hold between explanans and explanandum, and crucially whether the premisses of a demonstration convert. Those that are complete definitions clearly do—but Aristotle is not wedded to the view that every demonstration must contain at least one complete definition (it must contain at least one partial definition—a *per se* predication—but they do not convert). Consider a chain of explanation proceeding from genera to species:

180 if the species are taken individually, the explanandum has a wider extension than the subject: e.g. having external angles equal to four right angles extends further than either triangle or quadrilateral, but for all of them it has the same extension (as many as have external angles equal to four rights). And the middle term is an account of the first extreme, which is why all sciences come about through definition. Fox example, deciduousness applies to both vine and fig, and extends beyond each, but not beyond all of them [i.e. all the relevant species of tree], to which it is equal. Hence if you take the first middle term, it is the account of deciduousness. For there will be a middle term in the other direction, that all of them have some characteristic, and there is a middle of that too, that the sap coagulates, or something of the sort. What is deciduousness? The coagulation of sap at the seed-connection [i.e. the stalk]. (ibid. 2. 17. 99^a16–29)

If Aa B, and Ba C_1 , Ba C_2 ,. Ba C_n (where the C_i s are all species of some genus), then B explains why each of them is A; B is wider in extension than each of the species, but coextensive with the set of them. Suppose you have a scientific problem: why are vines deciduous? You may resolve it, Aristotle suggests (2. 16. 98^a35–$^b17$), by noting that vines are broad-leaved, and all broad-leaved trees are deciduous. This can be expressed syllogistically after the pattern of [C]; and Aristotle explicitly says in this case that being broad-leaved is the cause of deciduousness in vines, and indeed in anything else that is broad-leaved.

However, this is not the whole story: even though all broad-leaved trees are deciduous, the syllogism involving these properties and vines, namely

end p.163

[D] (i) all broad-leaved trees are deciduous

(ii) vines are broad-leaved

hence

(iii) vines are deciduous,

is not necessarily fully explanatory. After all, there is presumably some *reason* why broad-leaved trees shed. Their being broad-leaved does not, in and of itself, point to the cause of the shedding, at least without further explication. It is this which Aristotle hints at in the final sentences of **180**: broad-leaved trees are such that the sap in them coagulates at the nodes, which causes the leaves to snap off (presumably their heaviness is also a factor).

This suggests a further syllogism:

[E] ii(i) all sap-coagulators are deciduous

[E] i(ii) all broad-leaved trees are sap-coagulators

hence

[E] (iii) all broad-leaved trees are deciduous,

which in conjunction with [D] will provide the answer to the original problem; according to the earlier schematism, the closest middle term (B) to the explanandum is 'sap-coagulator'. Thus the 'middle term in the other direction' (i.e. of the subject) will be the earlier alleged explanatory middle term, i.e. broad-leaved—the characteristic which unites deciduous trees. But the proper account of deciduousness must involve sap-coagulation.

Thus the discovery of a truly explanatory syllogism is no easy matter. This example suggests[6] that Aristotle considers there to be two basic types of explanation which may be labelled *A* and *B*.[7] In an *A*-type explanation, you show that some predicate *P* attaches to a subject *S* because there is a middle term *M* such that all *S* are *M*, and *P* attaches of necessity to all *M*. Type-*A* explanations, then, need not pick out the genuine cause (in Aristotle's terms) of the attribute's holding of the subject. A *B*-type explanation, by contrast, explains why members of the class in question have the property they have: *S* are *P* in virtue of their being *M*. *M* is, in this case, coextensive with both *P* and the general class of individuals picked out by the original middle-term ('broad-leaved trees').
But here a further issue arises. Aristotle devotes some time in *Posterior Analytics* 2. 16–17 to discussing the question whether an explanans has to be coextensive with its explanandum: and his answer is, for genuine explanations at least, that it must (2. 17. 99^a1–17; cf. the opening sentences of
end p.164

180: Aristotle's account is, however, obscure and apparently contradictory). Yet he acknowledges that there are some cases in which what is apparently the same property will be given different explanations for different types of thing. Longevity is a case in point: it is caused in quadrupeds (at least so Aristotle alleges) by the lack of a gall-bladder, but in birds 'by dryness, or something else' (2. 17. 99^b4–7). Thus we may have distinct explanatory syllogisms *A*a *B, B*a *C* ⊢*A*a *C* and *A*a *D, D*a *E* ⊢*A*a *E*, where *C* and *E* represent different species, *B* and *D* distinct middle terms (2. 16. 98^b25–33). This has implications for the treatment of division and differentiae, to which I turn a little later on (one might choose to treat 'longevity' as being in some sense ambiguous, or rather as not denoting a single property, simply because what it was to be long-lived for a quadruped was different from what it was to be long-lived for a bird: cf. III.1*b*).
But at least in the case of *B*-type explanations, the middle term and the major (the explanandum) will be coextensive, since if it is wider in extension, the major premiss will be false, while if it is narrower, there will be some individuals comprehended under the major to which the alleged cause does not apply. In that case, either the major does not form a coherent class, or we have got the wrong explanation for it: 2. 16. 98^b33–8, 17. $99^a1–^b7$). Thus the middle term and the major will convert: *B*a *A* will be true as well as *A*a *B* (all and only sap-coagulators are deciduous). But in that case, we can construct valid syllogisms by a different route: we can prove that every vine is broad-leaved by showing that every vine is deciduous and all deciduous things are broad-leaved. In general, if *A* and *B* convert, then, supposing the minor *B*a *C*, we may infer *A*a *C*; but taking that conclusion with the converted major premiss *B*a *A*, we may equally deduce *B*a *C*.
It is plain that these syllogisms cannot both be explanatory, or at least not in the same way (cf. *Physics* 2. 3. 195^a9–11: a man is fit because of his training, and training because of his fitness—but in the one case efficiently, in the other finally). But how can we decide between them? Plainly this problem is central to the coherence of his view of cause and explanation:
181 knowledge of the fact and knowledge of the reason differ, primarily in the same science, and then in two ways. First if the deduction does not come about through

immediate premisses (for then the primary cause is not contained in them, and knowledge of the reason depends on the primary cause); and secondly if through immediates, but not through the cause but the more intelligible[8] of the convertible terms. For nothing prevents the non- explanatory of the convertible terms at times from being the more intelligible, so that the proof will proceed by way of it, e.g. the proof that planets are near
end p.165

because they do not twinkle. Let *C* stand for planets, *B* for not twinkling, *A* for being near. Then it is true to say *B* of *C*, since the planets do not twinkle. But also to say *A* of *B*, since what does not twinkle is near (this may be established either by enumeration of cases, or by perception). Thus *A* must belong to *C*, and so it has been proved that the planets are near. But this is a deduction not of the reason but of the fact, since it is not because the planets do not twinkle that they are near, but because they are near that they do not twinkle. (*Posterior Analytics* 1. 13. 78^{a}22–39)

Facts that require explanations are generally more familiar (to us at least: n. 1 above) than their explanations; and they can figure in sound deductions. Aristotle's example is instructive: all and only near bodies are non-twinklers (hence both *A*a *B* and *B*a *A* hold). Consequently, we can create two syllogisms, one of which concludes that the planets are near (because they don't twinkle), the other that they do not twinkle because they are near. Each inference can serve a purpose (we may indeed learn that the planets are near by way of an inference from their steady glow): but only the second gives us a reason why, and hence a scientific demonstration.

Aristotle offers another example, in which the terms are 'moon', 'spherical', and 'phase-exhibitor' (1. 13. 78^{a}29–^{b}12). His analysis is in many ways attractive: we surely do believe that it is the moon's sphericity (along with other relevant conditions, such as the fact that it shines by reflected light) that accounts for its phases, and not the other way around; and equally it would be bizarre to suggest that the planets' non-twinkling status *causes* them to be near. Again, we take the interposition of the earth to cause the eclipse of the moon, and not the other way around (2. 16. 98^{b}19–24). The reason for this is that we assume that phenomenal properties such as phase-exhibition, steady shining, and obfuscation rest upon, and hence are to be explained by, other ontologically more basic facts.

The extent to which we can determine the proper direction of explanation will depend upon the success with which we can show that one feature is in fact causally dependent upon the other; and this in turn may well be suggested by facts about how we can intervene to alter the situation in question. It would be absurd for me to try directly to make a planet to twinkle, and thereby cause it to move further away; on the other hand, if I want to make it twinkle, it would be sensible for me to remove it as far as possible from my vantage-point (we might seek in a similar vein to determine which of, for instance, broad-leavedness and sap-coagulation was the causally fundamental property). Modern accounts which eschew the notion of cause as an anachronism to be replaced by directionless functional relations (classically Russell 1912) are incapable of answering to these facts; equally, the Hempelian nomological-deductive model of explanation, where an instance is explained when it is shown to be
end p.166

subsumed under a covering law, is powerless to differentiate between such cases (since either syllogism, suitably recast, would be equally and indifferently explanatory).[9] If this is genuinely a deficiency (as I believe that it is), then there is indeed something to Aristotle's strictures here: and the dominant contemporary model of scientific explanation is inadequate.[10]

(b) The Epistemology of First Principles

Definitions are, as we have seen, fundamental to the project of exhibiting science in its appropriate, explanatory form. But how are those definitions to be arrived at? Science, and hence scientific explanation, is for Aristotle to be construed realistically. Our definitions cannot merely be convenient taxonomic stipulations: rather they must take us to the causal heart of the way things actually are. But this poses immediate epistemological difficul- ties. As is by now obvious, Greek natural scientists disagreed violently among themselves about the ultimate structure of reality, and their particular theories took off from starting-points which were anything but universally agreed upon (this endemic dispute was to be fastened upon by the Sceptics: VIII.2).
Moreover, as Aristotle is himself at pains to point out (*Posterior Analytics* 2. 1–8), definitions cannot themselves be demonstrated, since they are fundamental and hence fulfil criteria (*a–f*) of **179**. He needs some other secure method for establishing the elements of science and discovering the basic structure of reality. The details of Aristotle's complex and convoluted answers to these pressing questions are beyond the immediate remit of this study: but I shall briefly sketch one interpretation of them. Plato describes the method used to ascend to his first principles, knowledge of the Forms, as dialectic (III.3*b*, **134–5**). Aristotle also uses the term, but for him it has a more precise meaning: dialectic is the art of reasoning from agreed premisses (*Topics* 1. 1: the whole of the *Topics* consists in an investigation of dialectic); and some have thought[11] that Aristotelian dialectic is also intended to provide the foundation for the sciences. This is not, however, what Aristotle actually says:

182 deduction is an argument in which, when certain things are posited something else follows of necessity from what has been assumed. It is a demonstration whenever the deduction is from true and primary premisses, or from those which are such that we have grasped our original knowledge of them from true and primary premisses. A deduction is dialectical when it deduces from generally accepted beliefs (*endoxa*). True and primary things command

credibility not by way of others but through themselves: for it is not necessary to seek the why in the case of the scientific first principles, but each of the first principles is credible in and of itself. (*Topics* 1. 1. $100^{a}25–^{b}22$)

Demonstrative first principles are, then, in a sense, self-evident. But self- evidence cannot amount to obviousness, otherwise there would be no disagreements about them. In the final chapter of *Posterior Analytics*, Aristotle sketches an answer to the question of how we may arrive at these elusive, if evident, principles.

The method by which we arrive at first principles is called by Aristotle *epagōgē* (usually—if misleadingly—translated 'induction'), the process of arriving at universals from particulars. That process is outlined (although not explained) in our chapter. Starting from individual perceptions of things, the perceiver gradually, by way of memory, builds up an 'experience' (*empeiria*), which is 'the universal in the soul, the one corresponding to the many' (*Posterior Analytics* 2. 19. 100^{a}6–8); and it is this which provides the *archē*, or first principle:

183 these dispositions are not determinate and innate, nor do they arise from other more knowledgeable dispositions, but rather from perception, just as when a retreat takes place in battle, if one person makes a stand, another will too, and so on until the *archē* has been attained. (2. 19. 100^{a}9–13)

This process gives us universals (such as 'man'), without which we cannot utter assertoric sentences, which in turn lead to higher-order universals, such as 'animal' from particular species (2. 19. 100^{a}15–^{b}3). It is described in causal, not inferential terms (which is why 'induction' is misleading):[12] the world simply impresses us in such a way that we come to internalize ever wider and more inclusive concepts. We are by nature equipped to take on form in this way: if we are diligent and unimpaired, our natural faculties will see to it that we do so. Thus in a relatively literal sense we just come to see that Callias is a man, and, ultimately by the same process, what it is to be a man.

There is much that is sketchy and obscure in this account. But Aristotle is at least in a position to answer in principle the most pressing questions surrounding his scientific epistemology. Since the process by which we internalize form is not inferential, there is no regressive difficulty about specifying what previous knowledge is required in order to generate the resulting understanding of principles. The process gives us understanding of concepts and their extensions; but that in turn allows us propositional, definitional knowledge. For, as Aristotle stresses earlier, to know what a thing is is to know its definition (2. 1).

end p.168

(c) Division, Definition, and the Biological Sciences

In the opening chapters of *Posterior Analytics* 2, Aristotle argues that a definition cannot be demonstrated (2. 4), nor can it be proved by division (2. 5). In the methodological proem to his *Parts of Animals*, Aristotle equally argues that the method of division is not a means to the isolating of animal species. Aristotle was deeply interested in the comparative structures of animals, and in their inter-species relationships, and while it is disputed whether he was concerned with taxonomy in anything like the Linnaean sense,[13] he clearly is at pains to isolate and account for peculiarities, cases which seem not to conform to taxonomic type. The dissection of the blind mole to reveal its vestigial eyes (*History of Animals* 4. 8. 533^{a}1–11) is described not as part of a complete account of the

mole's structure, but rather to demonstrate that this singular instance of a viviparous quad-ruped lacking one of the normal senses is caused by an imperfection in the working-out of nature: the mole is not, as regards its basic form at least, deficient.[14]
Aristotle's purpose in dividing animals into classes is to exhibit what structures (along with their associated functions and activities) typically go together, and ultimately to isolate the causes of the variousness of animal structure thereby. In the course of this analysis, Aristotle develops and deploys a complex system of categorizing various degrees of structural similarity. Things may be identical in type (such as one man's nose and another's: ibid. 1. 1. $486^{a}16$–19); they may be morphologically similar, as in the case of the parts of different species of the same genus (such as different bird species' beaks: 1. 1. $486^{a}20$–$^{b}17$); or they may exhibit the even more remote relationship of analogy, as hand to claw, hair to feather and scale, etc. (1. 1. $486^{b}17$–21; cf. 3. 9–10).[15]
What matters for Aristotle is not the differences between animal species as such, but rather those differences which point to their causal relations with one another (these include whether the creature in question is blooded or bloodless, viviparous or oviparous, footless or footed). But Aristotle is at pains to explain (in *Parts of Animals* 1. 2–3) that simply dividing creatures according to certain differential characteristics provides no proper solution to the question of what their essential characteristics are. His target here is Plato's practice of dichotomous division (most
end p.169

completely, and tediously, expounded in the *Sophist*), in which by going down a branching tree of successive bipartite divisions we hope to arrive at the complete classification of the species in which we are interested. But in division of this sort, one branch of each bifurcation will generally be characterized in negative terms: and for Aristotle there can be no specific forms of negations such as featherless or footless. Animals falling into these categories do so by exclusion only, and not because they form a proper, coherent class (*Parts of Animals* 1. 3. $642^{b}22$–5). Proper division, on the other hand, ought to isolate genuine kinds; and that will frequently involve making more than one cut at the same level.
Moreover, some divisions, although not involving privative terms, mistake the actual order of things: thus people who divide animals into terrestrial and aquatic alone must artificially distribute the bird kingdom between the two groups (1. 2. $642^{b}9$–13). Equally to divide according to the attributes winged and wingless will artificially carve up such families as ants (1. 3. $642^{b}32$, $643^{b}1$). The appropriate method is to start with commonly recognized groups, such as birds and fishes (1. 3. $643^{b}9$–12), and see what combination of characteristics or differentiae sets them apart from one another.
The aim is ultimately to arrive at a definition of the species which states its essence by spelling out the genus (or larger kind) to which it belongs, plus the specific differentia which marks it off securely from the others. But, as we have already noted, definition is not simply a matter of marking the extension of a particular class, since that can be done by way of non-essential, albeit proper and necessary, attributes (§1*b* above). Rather definitions specify those properties which make things what they are, and which are themselves responsible for the other properties. The purpose of division is to arrive at

explanatory classifications of the differing animals which will explain why they possess the derivative properties they have.[16]

According to the methodology of *Parts of Animals* 1, division must proceed by way of progressive subdivisions of differentiae: you cannot divide into feathered and featherless, and then subdivide the feathered class into tame and wild, for tameness is not a type of featheriness (1. 3. 643^{b}18–22). A sound division will divide first into a general class (‘footed’) and then into a proper subdivision of it (‘two-footed’ or ‘cleft-footed’: 1. 3. 643^{b}27–644^{a}11). Here, ‘two-footed’ and ‘cleft-footed’ ‘cross-divide’: not all two-footed creatures are cleft-footed, and vice versa; consequently, what makes an animal two-footed cannot be what makes it cleft-footed, nor the other way round. The two traits are not causally linked, and if they are to have explanations, those explanations must reside in other facts. Indeed in

end p.170

Parts of Animals 1. 2–3 Aristotle comes close to rejecting division as a means of arriving at satisfactory definitions altogether; the final definition of an animal kind will, he thinks, generally involve specifying a variety of distinct differentiae arrived at by distinct diaeretic procedures, no one of which is necessarily more basic than or prior to any of the others.

And that last fact should come as no surprise: given his epistemology (§ *b* above), Aristotle cannot think that division, in and of itself, can supply us with definitions (and he explicitly rules out such a procedure in *Posterior Analytics* 2. 5). Rather, we must already in some sense know what it is that we have to define: we can just see that a certain set of creatures (swans, let us say) form a natural class. We may then proceed to investigate, on the basis of comparative study, just what it is that sets them apart, in terms of their causally basic properties. We may then, if we wish, exhibit the results in the form of a division (provided the strictures already mentioned are adhered to). But the division does not produce the definition; the dependence is rather the other way around.

2. Observation and Science

(a) Error and Empiricism

Aristotle regularly extols the virtues of observation. Biological research may not treat of the divine beings, but while inquiry into the latter is hampered by lack of evidence,

> **184** concerning perishable things—plants and animals—we are much better placed, since we live among them; and anyone who wants to take sufficient trouble may grasp many things about every kind that exists. (*Parts of Animals* 1. 5. 644^{b}28–32)

We should not shrink from empirical biology simply because some creatures seem loathsome to us (1. 5. 645^{a}15–31). He castigates Democritus' view that all mammals shed their milk-teeth since they are formed prematurely as a result of suckling; as Aristotle observes, some mammals retain their original teeth:

185 Democritus erred in his generalization about this through not having examined what happens in all cases; but it is essential to do this. (*Generation of Animals* 5. 8. 788^{b}10–19)[17]

He holds that it is apparent to both perception and reason that the heart is the first organ to develop in the embryo (2. 5. 741^{b}15–21), and he reports detailed and systematic observations of development in chicken embryos

(*History of Animals* 6. 3. 561^{a}6–562^{a}20). It is clear a priori that some organ must develop first, which is subsequently responsible for the creature's further development: its own internal principle of change and growth must appear as soon as the semen has induced its distinctive form-producing motions into the menstrual fluid, for only thus can the animal function independently. But that this first directing principle is the heart is learned from observation; and in general theory must answer to the empirical data (*Generation of Animals* 3. 10. 760^{b}28–33). The proper procedure, he states at *Parts of Animals* 1. 1. 639^{b}8–640^{a}18, is to begin by stating the appearances and only then to move from them to their causes.

Yet Aristotle is often accused of an excessive reliance upon arbitrary a priori assumptions, such as the superiority of right to left, front to back, top to bottom (*Progression of Animals* 4–5. 705^{a}26–707^{a}5; *On the Heavens* 2. 2: IV.3*e*), male to female. Females fulfil the role they do in generation since, lacking the male's innate heat and consequent ability to elaborate blood into dynamic, life-inducing semen, the best they can do is supply the matter for reproduction. Such commitments often led him astray: Aristotle was driven to postulate moon-dwelling animals, so that there could be creatures whose natural element was fire, to parallel those of the earth, water, and air (*Generation of Animals* 3. 11. 761^{b}16–24).[18] Furthermore, the observations which are supposed to support the theories themselves are sometimes suspect: he thinks, quite falsely, that it is ‘easily observable’ that in insect copulation the female inserts a part into the male (*History of Animals* 5. 8. 542^{a}1–17); and he supposes (quite why is unclear) that this ‘fact’ supports the view that the male contributes only form to generation.

Aristotle's observations are, in a sense, driven by the theories he espouses. He tried to observe and record the peculiarities of animal species in an objective manner; but he sometimes fails in his ideal, frequently lapsing into a rather charming credulity. For example, he claims that the ‘European bison’ (*bonasos*: it is not clear what animal this is supposed to be) defends itself by projecting its coruscating dung (‘which it can eject in prodigious quantities’: ibid. 9. 45. 630^{b}8–18) over large distances, as well as by constructing faecal fortifications. On the other hand, he can evince an admirable circumspection in drawing conclusions. While devoting much time to studying bees (5. 21–2, 9. 40. 623^{b}7–627^{b}22), he candidly admits that the data he has gathered are insufficient to prove his hunch that they generate asexually:

186 as far as theory (*logos*) and what seem to be the facts about them are concerned, this seems to be how things stand with regard to the generation of bees. But the facts have not been sufficiently ascertained, and if ever they are, then perception must be believed over theories, and theories only so long as they agree with the phenomena. (*Generation of Animals* 3. 10. 760^b28–33)

And some of Aristotle's more famous errors (such as the enormously influential theory of spontaneous generation) are based, albeit inadequately, on observation. In the cases where he held that species did not reproduce at all, he did so because he could discern no sexual organs, while he considered parthenogenesis (correctly) to be a possibility in other cases:

187 if there is a class which is female, but which has no separate male, it may generate an animal from itself. This has not yet been conclusively observed, but some cases in the class of fishes make us suppose this; thus of those called sea-perch, no male has yet been observed, although females with eggs have been [i.e. the species is not spontaneously generated]. But while we do not yet have conclusive evidence in these cases, there are also in the class of fishes those which are neither male nor female, such as eels. (ibid. 2. 5. 741^a33–^{b}1)

The observations upon which he based his notion that eels do not reproduce were accurate (eels do not develop reproductive organs before their migration to the Sargasso Sea); and his conclusion that they do not reproduce sexually was accepted until the nineteenth century. Impressive evidence of his observational skill may be found throughout the biology. He noticed the hectocotylization of one of the octopus's tentacles as well as guessing its function (*History of Animals* 4. 1. 524^a3–20); and he knew that some fish are externally viviparous, the embryo being attached to the parent by a cord (ibid. 6. 10. 565^b1–10). Neither of these observations was repeated until the last century. Aristotle far outstripped any of his Hippocratic and Presocratic predecessors in his pursuit of hard empirical data, amassing a vast repository of factual knowledge (not unleavened, it must be admitted, with mythical whimsicalities) about the natural world. But in spite of his own methodological protestations (*Meteorology* 1. 7. 344^a5–9; cf. *Generation of Animals* 5. 8. 788^b10–19, against Democritus), even he was susceptible on occasion to hasty over-generalization on the basis of inadequate evidence.

(b) Motion and the Elements

One reason why Aristotle thinks the Universe must be finite is that were it not we could make no sense of the natural motions of things. The concept of natural motion is entwined with those of place, weight, and lightness. Natural motions are proper to objects in virtue of their being the type of thing (elementally) they are. All the elements (earth, water, air, and fire) have an innate drive to move towards their natural places, which are

determined absolutely: the heavy seek the centre of the cosmos (which coincides with the centre of the earth: *On the Heavens* 2. 14. 296^b7–13), while the light tend towards the extremities (determined by the innermost of the heavenly spheres); and the places themselves possess powers (*Physics* 4. 1. $208^{b}11$). Thus the natural place for an element will be that region of space adjacent to the defining position which is capable of containing all of the element in question: in the case of earth, a sphere centred on the centre of the cosmos.

The elements will, other things being equal, occupy their natural places; and they will fail to move to them only if they are prevented from so doing, although this latter consideration shows how in a sense their particular motions are caused by things external to them, namely by the removal of obstacles to motion (*Physics* 8. 4. $255^{b}134$–$256^{a}3$; cf. X.4*a*). Lightness is not simply the privation of weight: both weight and lightness are positive, if contrary, properties (*On the Heavens* 3. 1–2, 4. 1–4), since what defines fire, for instance, cannot be negative, simply that which lacks weight. Aristotle contends that if fire is forced, as the least massive object, upwards by extrusion (as the Atomists say: VI.1*c*) it would still have weight: but then the greater the mass of fire, the more slowly it should rise, which is contrary, Aristotle thinks, to the phenomena (ibid. 1. 8. $277^{b}1$–9). Natural motions must be simple. There are generically only two kinds of simple motion, rectilinear and curvilinear (ibid. 1. 2. $268^{b}11$–18), of which the former has two proper species, up and down. This yields three possible basic motions, to which Aristotle assigns his three most basic elements: ether for the circular motion of the heavenly bodies, fire for upwards motion, and earth for downwards motion. Elsewhere he seeks (with dubious success) to make conceptual room for water and air, the remaining two of the canonical four sublunary elements, as being intermediate between earth and fire (cf. Plato, *Timaeus* 31b–32b: III.4*b*).

But natural rectilinear motion must, for Aristotle, be potentially completable, that is to say there must be some point at which the motion naturally comes to an end, which is where the object in question is naturally supposed to be: its natural place. Thus the idea of natural rectilinear motion requires that of natural place; and natural place can only be made sense of in a finite universe, since there can be neither centre nor extremity to an infinite space (*Physics* 3. 5. $205^{a}19$–20). Moreover, an infinite universe must contain an infinity of stuff (otherwise stuff would be infinitely rarely spread out through the cosmos, and hence could not form worlds); but then an infinite body could not have a natural motion, since there can be no differentiation between heavy and light in such a universe, and no proper distinction of direction (ibid. $205^{b}24$–$206^{a}7$).

In *On the Heavens* 1. 5–6, Aristotle applies similar considerations to the case of the heavenly bodies. These, located above the sublunary world, are

end p.174

subject to neither change nor decay (consequently, there is no real distinction of matter and form in them); and they move (at least according to *On the Heavens*) in virtue of a natural principle of circular motion, appropriate to their special fifth element, the ether. Since circular motion is (in a sense) infinite, there is no place where they naturally come to rest, and so they continue uninterruptedly for ever. But there can be no such thing as a circle of infinite radius (1. 5. $271^{b}26$–$272^{a}7$). Thus the heavens are limited in extent (Aristotle does not consider the possibility that the Universe might be infinite in virtue of its consisting of a never-ending sequence of concentric circles).

Evidently the success of all of this turns on the intuitive acceptability of the notion of natural place; why should we believe in any such thing? Aristotle's answer is, I think, twofold. First of all, such a supposition seems plainly supported by empirical evidence: heavy things do fall, and, in so far as they are free to do so, fall on a vertical trajectory towards the centre of the earth. Similarly, fire rises, and if there is no wind, it rises straight upwards; while the same goes for bubbles of air in water, and so on. These observations lead naturally to the conclusion that, unless they are interfered with by some external force, natural objects will always move in accordance with these internal tendencies (thus the theory embodies some notion of the conceptual idealization essential to all general scientific theories).

But in addition to its empirical support, the theory itself is elegant (at least at first sight), and general. It promises to account for a wide variety of phenomena on the basis of a small stock of primitive concepts. Moreover, it involves no action at a distance of the sort which made Newton's Universal Gravitation so difficult to swallow. This underlines another aspect of Aristotle's scientific explanations: they are commonsensical. He tries to explain the general structure and functioning of the world in terms of processes whose operations are evident to all.

(c) Dynamics and the Explanation of Motion

Equally commonsensical is Aristotle's treatment of dynamics:[19]

> **188** we see the same weight or body moving faster than another for two reasons, either because there is a difference in what it moves through . . . or because, other things being equal, the moving body differs from the other owing to excess of weight or lightness. . .. *A* then will move through *B* in time *C*, and through *D* (which is less dense) in time *E* (if the size of *B* equals that of *D*) in proportion to the density of the impeding body. (*Physics* 4. 8. 215^a24–$^b4$)

end p.175

Aristotle tries to develop an account of the proportionalities involved: an object will travel twice as fast through the same medium as one half its weight (7. 5. 249^b29–250^a16). His theory is not mathematized (he attempts no precise specification of the physical quantities involved); but it does at least try to bring the observable phenomena within the compass of general principles. Indeed, it is the theory's very faithfulness to the apparent phenomena that is its downfall: Aristotle is unable to attain to the degree of conceptual idealization required, for instance, to formulate Galileo's law of falling bodies. But it is one thing to account for the relations that hold between force, velocity, and mass; it is another to explain the mechanisms by which that force is imparted, and by which the motion of objects is maintained. And because Aristotle is unable conceptually to separate the notion of force from the substance in which it inheres, he is forced to the view that all agency is by contact (cf. *Physics* 7. 2, where all species of motion are reduced to pushing and pulling; and see XII.3*a*).[20] Furthermore, all motions must originate in some agent's natural action. A rod may move a stone, and a man's hand move the rod; but the man's

hand is moved by the man, and nothing moves him: he is the initiator of the motion (ibid. 8. 5. 256a6–33). The proximate and ultimate causes of the stone's motion are both causes in a way: but only the latter is properly the cause. Since the chain of action is continuous here, Aristotle has no difficulty in principle in accounting for its transmission.
However, projectile motion, where the activity continues after the agent has apparently stopped acting, poses problems. Projectile motion is, of course, forced, and so cannot (on Aristotle's general principles) be continuous and eternal. The fact that such motions ultimately come to an end is, then, no embarrassment for him. Rather the problem is that

> **189** if everything that moves is moved by something, how can some things which are not self-movers yet continue to move when the mover is not in contact with them, such as projectiles? (8. 10. 266b29–31)

It cannot be that the agent also moves something else (the surrounding air), which continues to impart motion to the projectile, since this simply pushes the problem a stage back: the air will still be moving after the thrower's hand stops.

> **190** We must thus say that the first mover makes the air (or water, or whatever else can both move and be moved) such as to be able to cause motion, but that this does not cease to cause motion and to be moved at the same time, but stops being moved as soon as what moves it ceases moving, but still continues to cause motion, and so it moves something else consecutive with it. And the

end p.176

> same thing is true for the latter. It stops in cases where the motive force engendered in the consecutive object is always less, and it comes to an end when the former can no longer make the latter a mover, but only moved. (8. 10. 267a2–11; cf. *On the Heavens* 3. 2. 301b18–30; *On Dreams* 2. 459a28–b1)

Successive parcels of air are endowed with the capacity to maintain movement in the missile, but with decreasing force. This unjustly derided solution is ingenious. Aristotle does not say that each successive parcel is made into a *self*-mover (contrary to the interpretation of Alexander of Aphrodisias: Simplicius, *On the ‘Physics’* 8. 10. 1346. 37–1347. 37), which would generate other difficulties: self-movers are the origins of their own motion and rest, so if something is made into a self-mover, it must continue to possess this motion-causing capacity, as a human being does, at least as long as it remains substantially unchanged.[21]
Even so, what stands in the way of Philoponus' much later view that the impetus is impressed directly into the projectile (XII.3*c*)? Aristotle may have thought that the material of the projectile was not of the appropriate type to be a motion-causer, unlike air and water. Why should air be a more suitable motion-causer than the projectile itself? Perhaps simply because we can observe, in other circumstances, the motive properties of air. Not every body can be directly imbued with force; rather some of them possess a natural capacity to receive the motion-causing power, which they can then impart to others.

(d) Sublunary Phenomena

Still, Aristotle needs to say more about how he conceives of the transmission of such powers in the natural world. He describes the subject-matter of his 'meteorology' as

191 everything which occurs naturally, but in a less orderly manner than that of the first element [i.e. ether] of bodies, in the place bordering most closely on that of the motion of the stars, e.g. the Milky Way, comets, meteors, everything we assign as being common attributes of air and water, as well as the kinds, parts, and attributes of earth, from which we will examine the causes of winds and earthquakes, and whatever occurs as a result of their movements . . . as well as the fall of thunderbolts, typhoons, firewinds, and the other recurrent phenomena which occur to their bodies as a result of condensation. (*Meteorology* 1. 1. 338^b27–339^a6).

His concern is with the explanation of phenomena lying outside the perfect domain of the heavenly spheres. Most have to do with the interrelations of air and water, since most alteration takes place in those elements; and even earthquakes turn out to have an atmospheric explanation. What

distinguishes such processes from those of the heavens is their lack of apparent regularity (it is for this reason that Aristotle assigns cometary phenomena to the sublunary world: ibid. 1. 6. 342^b30–343^b8).

Ultimately these phenomena owe their motive power to the sun, whose seasonal variations in proximity to the earth as it moves along the ecliptic ultimately account (efficiently at least) for all of the processes of the sublunary world (including rainfall: cf. IV.3*d*, **168** above):

192 the whole structure around the earth is composed of these bodies [i.e. water, air, fire]. . . . This is of necessity continuous with the higher motions, which control all of its power from there. For whence the origin of motion for everything comes should be considered as the primary cause. . . . Consequently fire, earth, and their congeners must be considered to be the causes of what occurs in this region, in the material sense: for this is what we call the substrate and patient, while the cause in the sense of the origin of motion should be ascribed to the power of the eternal movers. (1. 2. 339^a19–32; cf. **198–9**)

But we may none the less examine and analyse the sublunary mechanisms by which this basic source of power is translated into the multitude of terrestrial phenomena.

Aristotle's fundamental meteorological mechanisms are two 'exhalations', hot and dry, and hot and moist, which arise from the earth and nourish respectively the upper and lower parts of the atmosphere, the air and the fire-sphere (ibid. 1. 3). The fire-sphere is not literally composed of fire, 'since fire is an excess of heat and a kind of boiling' (1. 3. 340^b23): it is the inflammable part of the upper atmosphere, which is moved by contact with the innermost celestial sphere (1. 3. 341^a2). There is a constant interchange (the mechanics of which are left obscure) between the air- and fire- spheres (1. 3. 341^a7–9). The sun generates heat not because it is itself naturally hot (not being made of fire), but rather by friction (1. 3. 341^a13–37): 'moving bodies are often seen to melt', and 'the air about something being moved by force becomes particularly hot'.

Primitive observations are thus invoked to buttress the theory's plausibility; but they do no more than that. And we need such explanations of the sun's heating power only if we are already convinced it cannot be made of fire (as Aristotle is, on the basis of his general theory of the composition of the Universe). In any case, Aristotle's observations certainly do not entail his account: at most they weakly suggest it. But he is well aware of this:

193 we consider that we have given an adequate demonstrative account of things unavailable to sensation when our account is consistent with what is possible. (1. 7. $344^{a}5$–7)

That methodological remark prefaces his account of comets: the 'hot and dry exhalation' (i.e. the fire-sphere) is prone to ignite as a result of its

motion, causing meteors when the combustion is rapid, and comets when the burning is steady and slow (1. 7. $344^{a}13$–20). Meteors are compared with flame running through a thread of chaff (25–9), while

194 if this fire were not to run through and exhaust the fuel, but were to stand at the point where the inflammable material were densest, then the end of its progress would be the beginning of its orbit. (1. 7. $344^{a}29$–32)

Everyday experience is ingeniously (although non-probatively) invoked in order to demystify a remote phenomenon (cf. I.2*h*).

Aristotle's exhalation theory exhibits familiar strengths and weaknesses. He is sensible of the need to make the theory accord with empirical data, and of the way in which observations may be urged in its support. Some of these are simply false: comets are not, as he claims they are, more frequent in windy and dry weather (a 'fact' which is supposed to confirm their link with the hot and the dry: 1. 7. $344^{b}19$–25). But in some cases empirical observation does apparently confirm the theory: vapours are indeed seen to rise from the earth and the sea. Moreover, the theory is economical and general, capable of unifying apparently disparate phenomena, and of extension into domains that are at first sight remote: earthquakes are caused by exhalations trapped within the earth (2. 8; those that escape easily become winds). Aristotle knows that some places are more earthquake-prone than others, and suggests, reasonably enough, that geological variations are responsible for this discrepancy.

So far, the 'meteorological' explanations we have considered have been largely mistaken. By contrast, Aristotle's account of the origins, nature, and dynamics of the sea (2. 2–3) is a good deal more successful. It is reasonable, in view of its bulk and stability, to suppose the sea to be the origin of all water. Furthermore, it envelops the earth (albeit imperfectly) like the other spheres (2. 2. $354^{b}3$–15);

195 but another difficulty stands in the way of this view: if this body is the source of all water, why is it not drinkable but salt? The reason for this will constitute the resolution of this difficulty, and will ensure that our basic assumptions regarding the sea are correct. (2. 2. $354^{b}19$–23)

The sun's movement periodically draws up 'the finest and sweetest' (cf. II.1*c*, **83**) part of the water, leaving behind the grosser brine (2. 2. $354^{b}28$–33, $355^{a}33$–5, $^{b}19$); the system is self-contained, evaporation from the sea precisely balancing its replacement from rivers, even if it is less perceptible since it takes place across the entire surface of the water (2. 2. 355^b21–33).

Since the Universe is eternal, the sea must always have been there. Consequently, the world cannot be gradually drying out (as Anaximander held), and its hydrological system must be stable (2. 3. 356^b4–357^a3). Moreover ‘poetical’ accounts of its saltiness like Empedocles' that it is ‘the

sweat of the earth’ are ‘not conducive to understanding’ (2. 3. 357^a25–8). Aristotle's explanation invokes the dry exhalation (2. 3. 357^b24–6) and the general contention that residues, of which it is one, are salty (2. 3. 358^a3–12). The dry exhalation mixes with the moist in clouds, subsequently falling as salt-impregnated rain (2. 3. 358^a16–25). And this process ensures that the sea is continually becoming more salty, since the salt part is left behind, a fact which Aristotle claims to have observed experimentally (2. 3. 358^b16–18):[22]

196 that saltiness consists in a mixture of something is clear not only from what has been said, but also [from the fact that] if someone places a jar made of wax in the sea, having bound its mouth so as to prevent the sea getting in, the water which percolates through the wax walls becomes fresh: for the earthy part which causes the saltiness in the mixture is separated off as though by a filter. This is also the cause of its weight (for brine weighs more than fresh water), and of its density; for their densities vary so much that vessels laden with the same weight almost sink in rivers but ride well at sea and are sea-worthy. . . . An indication that the mass of a mixed substance is denser is that whenever one makes water highly salty by mixing salt with it, eggs will float on it. (2. 3. 358^b34–359^a14)

That passage is absolutely characteristic. It makes acute observations, and sketches two experiments. Each of them is well conceived to be such as to prove what it is supposed to prove; and the claims about relative densities are of course correct. But nothing can function as Aristotle alleges his wax jar (whatever that may be) does to desalinate water, since brine is not a mixture in the sense Aristotle suggests it is at all. And that brings us conveniently to Aristotle's chemistry.

(e) Chemical Combination

Aristotle's theory of chemical combination, which is outlined in *On Generation and Destruction* 2 and *Meteorology* 4, represents a considerable advance upon the traditional four-element physical chemistry he elsewhere adopts.[23] Four qualities, hot, cold, wet, and dry, are now primary (*On Generation and Destruction* 2. 2. 329^b17–330^a29), the four elements being composed from them (earth = cold/dry, water = cold/wet, air = hot/ wet, fire = hot/dry: 2. 3). Aristotle's theory is highly reductive: these primary ‘powers’, as Aristotle sometimes prefers to call them (*Parts of Animals* 2. 1. 646^a14–15), are themselves the foundation of further derived properties, such as heaviness and lightness,

condensation and rarefaction, roughness and smoothness (2. 1. 646a18–21), as well as malleability, ductility, and brittleness (*Meteorology* 4. 8–9).
end p.180

Aristotle distinguishes hot and cold from wet and dry, in that the former are principles actively involved in change, while the latter merely form the passive substrate for change (ibid. 4. 1. 378b10–27). The elements compose the first level of actual ontological existence (there being no such thing as the hot in separation), although the actual physical stuff we call earth is itself a mixture of elemental earth and water (*On Generation and Destruction* 2. 3. 330b22–30). Cold solidifies and compacts objects, while heat solidifies what is uniform and dissolves what is discontinuous.

This account is clearly less than strictly empirical; yet Aristotle's willingness to abandon the strict a priori parallelisms and contrarieties of his predecessors shows him striving to accommodate his theory to the phenomena: heat does indeed solidify some things (e.g. clay: *Meteorology* 4. 7. 384b19–21, 9. 385b9). His theory seeks to explain the various distinctive chemical and physical changes in terms of fundamental, elemental properties, whether singly or in combination; and it develops a clear conception of the difference between real combination (which occurs in the formation of uniform parts) and mere juxtaposition or mixture (*On Generation and Destruction* 1. 10, 2. 7; cf. *On the Senses* 3. 440a31–b25), albeit one quite different from our own.

(f) The Structure of the Universe

Aristotle conceived of the Universe as a sequence of concentric spheres, the outermost, which contained the heavenly bodies, being composed of the incorruptible, immutable, rotating element ether (*On the Heavens* 1. 2–4). The mathematics of his astronomical system was fundamentally that of Eudoxus, a younger contemporary of Plato's, who devised a system of concentric spherical motions to meet Plato's challenge and reduce to order the ‘intractable and marvellously intricate’ wanderings of the planets (III.5, **153** above):

197 Eudoxus of Cnidus was the first Greek to concern himself with this sort of hypothesis, since Plato, as Sosigenes says, had set it as a problem for all serious students of the subject to find what uniform and ordered motions are such that, being assumed, the appearances in the case of the planetary motions may be saved. (Simplicius, *On ‘On the Heavens’* 2. 12. 488. 18–42)

The system, which is of great ingenuity, merits a brief description here.

Eudoxus proposed to account for the various anomalies in observed planetary and lunar motion by a model in which distinct spherical motions were superimposed upon one another. Thus he took over from Plato (and no doubt the tradition) the idea of the sphere of the fixed stars rotating in one direction, and of motions for the bodies of the solar system in the opposite sense and at the inclination of the ecliptic. But in order to give
end p.181

something approximating to an accurate mathematical description of precisely how the planets (and the sun and moon) behaved, he introduced further, nested spheres, set at various angles to one another, and rotating in different directions, for each of the bodies of the solar system, the body itself being carried on the innermost sphere of the nest, and whose observed motion was thus the result of the sum of the various spherical motions. Eudoxus' work survives only in later, somewhat garbled reports, and the precise reconstruction of his scheme is controversial.[24] None the less, it is clear that it was designed primarily to account for the observed variations in latitude of the planets and the moon,[25] as well as the retrogradations of the outer planets.

It is in the case of the latter (already known to Plato: III.5) that Eudoxus' ingenuity is most apparent. In addition to the two basic spheres (the first for the diurnal motion of the celestial sphere, the second to represent the planet's periodic west–east path along the ecliptic) common to all the bodies in the system, he added two inner spheres. The first of these (the third sphere) was inclined at right angles to the second (i.e. its poles were to be found on the ecliptic), while the fourth, which carried the planet, was set at a slight angle to the third (the angle varying with each of the outer planets), revolving at precisely the same velocity but in the opposite direction. The geometrical result of this, Eudoxus was able to show, was that the planet described a sort of figure of eight around the ecliptic[26] as a result of the motion of the inner spheres, but a figure of eight dragged out in the direction of the periodic motion along the ecliptic by the second sphere. The resultant apparent motion produced both retrogressions and variations in latitude.

The precise results, however, were of varying accuracy: they give reasonably good values for Jupiter and Saturn, but were much less satisfactory in the case of Mars. None the less, for our purposes what matters is not so much the empirical adequacy of Eudoxus' construction,[27] but rather the striking manner in which it showed that it was at least plausible to suppose

that Plato's challenge could be met. Eudoxus' system was modified by a later contemporary, Callipus, by the addition of one extra sphere each for Mars, Venus, and Mercury, and two for the sun and the moon (Aristotle, *Metaphysics* 12. 8. 1078^{b}33–8), the two spheres for the sun being designed to account for the inequality of the seasons ignored in Eudoxus' scheme, those for the moon no doubt to attempt a better approximation to the moon's extremely complex latitudinal deviations, while the extra planetary spheres presumably sought better to account for their stations and retrogressions.[28]

These models were purely mathematical devices, designed to evince the underlying regularity of planetary motions. Although Aristotle took this over, he attempted in addition to give the structure a physical interpretation by postulating a set of counter-spheres interposed between the spheres of the heavenly bodies to balance their forward motions, and to explain how a sequence of nested spheres need not become progressively more and more rapid the further into the sequence one goes (this is described in most detail, albeit still sketchily, in *Metaphysics* 12. 8). The heavens are a system of interlocked spheres, moving like a vast astrolabe as a result of their inherent, natural, elemental nisus.[29]

(g) The Motor of the Universe

That at least is the view expressed in *On the Heavens*; and it appears to be perfectly self-contained. The celestial bodies move continuously, simply because it is the nature of the element which makes them up to behave like that; and that motion is also, derivatively, responsible for everything else that occurs in the sublunary regions: for even the generation of animals is seasonal, and the seasons are caused by the sun's movement along the ecliptic:

198 (i) furthermore, since it has been shown that motion (*kinēsis*) in the form of locomotion (*phora*) is continuous [**199–200** below], it is necessary, this being

end p.183

so, that generation be continuous as well, since the locomotion will produce generation uninterruptedly by means of the approach and retreat of the generator [i.e. the sun]. (ii) At the same time, it is clear that what was said earlier, namely that the primary change is locomotion and not generation, was well said: for it is far more reasonable for what is to be the cause of generation for what is not than for what is not to be the cause of the being of what is. (iii) So while what is moving is, what is being generated is not; for this reason locomotion is prior to generation. (iv) And since it has been both suggested and shown that both generation and destruction in things is continuous, and we say that locomotion is the cause of generation, it is clear that if locomotion is single, both of them cannot occur, since they are opposites. (v) For the same thing in the same state always naturally produces the same thing; and thus there would either always be generation or always destruction. (vi) The motions must be plural and opposite, either in locomotion, or in irregularity: for opposites are the cause of opposites. (*On Generation and Destruction* 2. 10. 336^{a}15–32)

Unless there is a continuous source of motion (or change: the word *kinēsis* for Aristotle covers more than simply local movement), Aristotle thinks, the processes of the world cannot themselves be continuous. He is of course positioning himself here squarely in the centre of a tradition that stretches back through Plato (III.2*a*, **131**) ultimately to Anaximander (I.2*d, f*, **6**, **8**). Moreover, **198**(ii–iii) show him sensitive to the Parmenidean restrictions on generation: continuous locomotion, which is an actuality, is to be responsible for the generation of things. (iv–vi) introduce further familiar considerations (effectively, Plato's P2: III.1*b*; cf. also P3 and P4). ‘Irregularity’ here is equivalent to eccentricity in the geometric sense. There is, in one sense, only one continuous circular motion for the sun: it does not actually change direction. But the inclination of the ecliptic has the effect of making the sun now approach and now retreat (ibid. 2. 10. 336^{a}32–^{b}10), and this ‘irregularity’ is enough to produce generation at one time and destruction at another, consistent with the basic causal principles.

That there can indeed be a continuous, circular form of locomotion (**198**(i)) is argued in the *Physics*:

199 we must examine whether there can be some continuous motion or not, and if there can, what it is, and which is the primary motion. For it is clear that if it is necessary that there always be motion, and if some motion is primary and continuous, then the Prime Mover (*to prōton kinoun*) will produce this motion, which is necessarily one and the same, continuous, and primary. (*Physics* 8. 3. 260^a20–6)

This is the Prime Mover's (PM) first appearance: and so far its role seems relatively innocent. Indeed, Aristotle might simply be referring to whatever it is which is, in virtue of its particular continuous locomotion, as a matter of fact the efficient cause of the never-ending sublunary cycles of

end p.184

generation and destruction (i.e. the sun). A later stretch of argument makes good on the promise to demonstrate the primacy of rotational locomotion over other forms of change:

200 consequently, since it is necessary that motion be (*a*) uninterrupted, and if uninterrupted then either (*b*) continuous or (*c*) successive, and rather continuous, since the continuous is better than the successive, and we always assume that the better obtains in nature if it is possible, and it is possible for it to be continuous . . . and this can be nothing other than locomotion, then it is necessary that locomotion be primary. For there is no necessity why what is moved should grow or alter or be generated or destroyed; but none of these things are possible without the continuous locomotion caused by the Prime Mover. (8. 3. 260^b20–9)

200 merits brief analysis. The last sentence explains the priority of locomotion: it is presupposed by the other types of change (cf. 8. 3. 260^a26–$^b19$, b29–261^a28, 9. 265^a13–266^a5), which is why 'motion' is appropriate to all of them. The source-motion must be (*a*) uninterrupted, since if it were not it would require some further, more basic source to set it going again—to which the same considerations would apply (cf. I.3*c*; III.2*a*). The only form such motion can take is for it to be either (*b*) continuous (i.e. for it to exhibit a constant and unchanging form), or (*c*) successive (i.e. for each form of motion to be succeeded immediately by some other, without end). But (*b*) is preferable to (*c*), since it would be better for things to be that way (cf. 8. 8). This appeal to the good is reminiscent of moves made elsewhere (IV.1*d*, **168**(i)), particularly in the biological teleology (IV.1*e*); but whereas in most of those cases the sense in which things are arranged for the sake of the good, or for the better, could be explicated in terms of contribution towards the overall well-being of the creature in question, here more abstract issues seem to be involved.

It is tempting to take the appeal to the better here as being simply part of a general methodological principle: other things being equal, it is a useful regulative heuristic to assume, if it is at all possible, that nature will have arranged things for the best, or have adopted the best (in the sense of the most perfect) methods. But there is more to it than just that. For Aristotle, natural explanations structured in accordance with the best are not merely better, in the sense of being more economical, or having greater explanatory potential (although of course they are): they are more likely to be *true*, in a strongly realistic sense. It is, for him, a fact of nature that it strives for perfection (compare his appeals in biology to the fact that a structure is arranged in the 'most honourable' fashion: IV.3*e*, **170–2**); and he seeks to explain the direction of heavenly rotation on similar grounds:

201 if nature always produces the best of the possibilities, and it is the case that, just as of the rectilinear motions that upwards is the more honourable (since

end p.185

the upper place is more divine than the lower), in the same way motion towards the front is more honourable than towards the rear, then it [i.e. the universe] too contains a prior and a posterior, given that it has a right and left, as was stated earlier. . . . For this is the explanation that solves our problem, since if things are the best they could be, this will be the explanation for what has been said: for it is best to move with a simple and unceasing motion, and that in the direction of the more honourable. (Aristotle, *On the Heavens* 2. 5. 288^{a}3–12; cf. 2. 2)

Perfection and intelligibility go together for Aristotle: and the world will be maximally intelligible if it is such as to be explained on the basis of a minimum of hypotheses. Thus, postulating one continuous motion is simpler than having a succession of them; and hence (*b*) of **200** is preferable to (*c*). Congruently, (*b*) leaves less unexplained: if (*c*) were in fact the operative model, we would further need to account for *how* each motion brought about its successor, and for what the fundamental structure of that succession might be; if there is no real answer to that last question, then the world is not maximally intelligible—but if there is, once more we have been reduced to an explanation of type (*b*).

This brings us to the Prime Mover. For all that has been said so far, the heavenly rotations might simply be a brute fact, to be explained (as *On the Heavens* has it) purely in terms of the natural motions of their elemental components. But for whatever reasons, Aristotle ultimately found this unsatisfactory, presumably because it still left something apparently to be explained: *why* should things behave in this way? If that question is unanswerable, then in some degree the world falls short of full intelligibility. Consequently, the heavenly bodies are elsewhere said to move in emulation of the PM, the highest and most divine being, whose example of perfect, uninterrupted contemplative activity inspires everything else in the universe to imitate it to the best of its ability (*Metaphysics* 12. 7. 1072^{b}3–4). Aristotle allows that all of the heavenly spheres which he introduces to give astronomy a physical interpretation are propelled by unmoved movers (12. 7. 1073^{a}26–1073^{b}1, 1074^{a}14–18), which are not (contrary to *On the Heavens*) to be identified with the spheres themselves; but these are not themselves Prime Movers:

202 there is something which is moved with an unceasing motion, namely that in a circle; and this clear not only from reason, but also in fact: consequently, the first heaven [i.e. the outermost sphere] will be eternal. But there is also something which moves it. And since what is both moved and moves is intermediate, there is some motor which causes movement but is not moved, is eternal, and is substance and actuality. And it moves as follows: objects of desire and thought cause movement without being moved. (12. 7. 1072^{a}21–7)

end p.186

Thus the PM is a cause of motion, but is itself motionless. Aristotle here abandons PCS (the idea that causes must resemble their effects: I.3*c*), at least in full generality. The reason, once again, is explanatory simplicity. The highest cause must be that which exists in pure, unadulterated actuality (otherwise further explanation will be needed to explain how it moves from potentiality to actuality: cf. 12. 6. $1071^{b}12$–23): yet even the heavenly bodies in a sense contain potentiality, in the same sense as they involve matter (they can move from place to place: 8. 1. $1042^{a}34$, $^{b}6$). Hence if they are to be basic, there will be something left without explanation, namely their ability to actualize their potentials (cf. 12. 7. $1072^{b}4$–12). So the PM cannot itself be in motion (*Physics* 8. 10. $267^{a}20$–$^{b}17$). But it is vital to realize that the PM is a cause only in so far as it is the object of desire. It does not directly impart motion to the spheres; rather it excites in them the desire to emulate, in so far as they are capable of doing so, its state of pure intellectual activity. Their own motion is, efficiently at least, self-caused (they are thus alive: *On the Heavens* 2. 2. $284^{a}32$–4, $285^{a}27$–30). But they rely upon the PM to provide them with a reference-point, as it were, something upon which to fix their gaze in their drive for as complete an actualization as possible; moreover, the structure given by the mere existence and activity of the PM to the rest of the cosmos stretches all the way down through the hierarchy of nature.[30] Everything strives, to the extent to which it is by nature able, to ape the perfection of the most divine being of all:

203 since we say that nature always desires the better in all cases, and it is better to be than not to be . . . while this cannot obtain in all cases on account of their distance from the first principle, God filled up the whole in the manner still available by making generation perpetual: for in this way being would be the most coherent, since eternal generation and destruction are closest to being. And the cause of this, as has frequently been stated, is the circular motion, for it is the only continuous one. For this reason, all the other things which change into one another in respect of both active and passive properties, such as the simple bodies, imitate the circular motion. For whenever air is generated from water, and fire from air, and then again water from fire, we say that the generation has gone round in a circle since it has returned to where it set out. Linear motion is thus continuous by its imitation of motion in a circle. (*On Generation and Destruction* 2. 10. $336^{b}27$–$337^{a}7$)

Different types of thing are capable of mimicking the PM to a greater or lesser extent—but they all do so in so far as the constraints imposed by their natures will allow (*On the Heavens* 2. 12. $292^{a}22$–$^{b}25$). Thus the PM is not an efficient cause (although later commentators disagreed about this:

end p.187

Simplicius, *On 'On the Heavens'* 271), even though Aristotle insists that its effects are felt primarily at the circumference of the heavens (which is where it is located: *Physics* 8. 10. $267^{a}20$–$^{b}9$); but by serving as the ultimate cosmic exemplar it gives structure and intelligibility to the world as a whole: indeed it is what makes the Aristotelian world a whole.

3. Theophrastus and After

(a) History, Science, and Method

When Aristotle died, his Lyceum was taken over by his long-time pupil and associate Theophrastus (*c*.370–287 bc). Theophrastus continued to further Aristotle's scientific and philosophical projects, filling in the gaps in his great account of the natural world with two texts on plants which survive: *History of Plants* and *Causes of Plants*. Although he wrote a brief *Metaphysics*, and an *On the Senses*, which also survive (as well as texts on logic and ethics, most of which do not), Theophrastus' main interest appears to have been in the physical world. He wrote treatises *On Fire* and *On Winds*, as well as a *Meteorology*, which survives in an extensive Arabic epitome.[31] A long report in Philo[32] (*On the Eternity of the World* 117–49 = Fr. 184 FHS&G) shows him to be an energetic and ingenious defender of the Aristotelian doctrine that the world has always existed against a variety of physically based objections.
Why, if it is eternal, is the earth's surface uneven? Erosion should have flattened everything out (118–19). Not so, Theophrastus replies, since mountains are continually being raised by volcanic action and other forces (132–7). Secondly, it is objected, the sea can be seen to be drying up—if this process had gone on for ever it would have been exhausted (120–3). Again Theophrastus counters that the recession of the sea in some areas is part of a cyclical process, and is compensated for by inundations elsewhere (138–42). The fact that technology is relatively recent (127–9) is accounted for by the fact that there have been periodic set-backs in human history (145–9). Most generally of all, Theophrastus responds to the argument that, since each of the elements is destructible, the whole which is composed of them must be destructible too (127–9), by noting that this involves a fallacy of composition—only if all elements could be annihilated at once would the universe thereby be destructible (133–4). Again, his refutation rests upon a notion of conservation, but one that is underwritten by the deep and abiding Greek aversion to creation *ex nihilo* and complete annihilation (CP1*a*, CP2: I.3*a, d, e*, **28**, **40**, **43**; III.4*a*; IV. 1*b*, **156–7**).
end p.188

But his most enduring contribution to the history of thought was perhaps his invention of the history of thought. Aristotle had of course been interested in and drawn upon the views of his predecessors; but Theophrastus is the true originator of the great Greek doxographical tradition, and much of what we know of early Greek philosophy (e.g. Xenophanes' theology: I.3*a*, **27–8**) derives, via the Aëtius doxography, from his work in the field. It emerges from the fragments of Theophrastus' doxographical works (including a *Natural Researches* and a *Physical Opinions*) that one of his principal concerns (following Aristotle's lead in *Physics* 1. 1–4) was with cataloguing and comparing the various accounts of the natural philosophers concerning the nature and number of explanatory principles to be postulated; many of these fragments are preserved in Simplicius&*On the 'Physics'*: 22–8, 149–50, 154.[33]

The surviving *On the Senses* devotes much space to dealing with earlier philosophers' views on perception. It is of only tangential relevance to our concerns; but Theophrastus' criticisms of his predecessors involve assessing their accounts relative to his own methodological principles of explanation. Thus, in the case of Empedocles (whose version of the 'like-perceives-like' theory Theophrastus deals with at particular length: *On the Senses* 12–24), a number of explanatory inadequacies and lacunae in the theory are exposed. The appropriate fit of effluences into pores cannot, Theophrastus thinks, *on its own* account for sensation (ibid. 12: unless we assume, for example, that every pipe feels what it contains). Empedocles distinguishes the internal fire in the eye from the external fire, but does not explain why the former should be susceptible of perception (13); he denies the existence of void, yet holds that the perceptual pores are sometimes empty (13); he holds that perception can be adversely affected by the clogging of the pores by foreign particles, and yet that only the perceptually appropriate particles can enter the pores (14); moreover, his theory conflates the 'like-perceives-like' notion with the idea that perception is by contact: yet these are distinct concepts, and hence Empedocles has offered two separate explanations for one phenomenon (15–16).

In the case of hearing, Empedocles holds that the internal reverberation hears the external one: but in virtue of what? There is no explanation given as to how the former can perceive but not the latter (21); while in the case of smell, Empedocles fails to give a unified explanation of the phenomenon, since he explains it on the basis of respiration, yet some animals smell without breathing: respiration is thus not in itself a cause of smelling, but only incidentally (21–2). Finally, Theophrastus accuses Empedocles of confusing an instrument with a real cause (cf. XI.4*b*): the composition of the blood cannot be the real cause of the bodily functions any more than

the tongue is the real cause of speaking well, or the hands of construction: rather it is the form that it embodies which is the real cause (24). These concerns are evident too in his treatments of other thinkers' views: for example, he chides Plato for failing to state the proper explanation for sensation (91), and Democritus for neglecting to explain how his perceptual images can become inverted in reflection (52), as well as for failing to give a common cause for all cases of smell (71).

Several general methodological principles of explanation may be extracted from these (and numerous other) telegraphic examples. Explanation should not proceed from one unknown to something equally unknown, or court regress; unified sets of phenomena should be given unified rather than disparate explanations; explanations should be consistent, both with themselves and with other aspects of the doctrine; they should be economical; they should not be categorially confused; and they should be complete.[34]

Many of these concerns (notably that of explanatory completeness) may be seen in action in the general account of the interrelations between plants offered in his botanical works. An instructive passage from *On the Causes of Plants* displays his awareness of the importance of conceptual refinement, abstraction, and idealization in science. Having assigned the cause of early budding and growth to the twin influences of external warmth of air and sun and the plants' internal temperaments of hot, cold, wet, and dry, he remarks:

204 while wetness and dryness are pretty much discernible by the senses, hot and cold, which involve reason rather than perception, are the source of disputes and controversies, as does everything subject to reason's adjudication. It is important to define these concepts, particularly in view of the fact that many things are explained by way of these principles. In fact, we must assess all concepts of this sort by their empirical consequences, for it is on the basis of facts that we arrive at conclusions regarding the nature of elementary forces. (Theophrastus, *On the Causes of Plants* 1. 21. 3)

The open acknowledgement of the fact that Hot and Cold are no longer purely perceptually determined principles, but rather theoretical terms that require explication, shows Theophrastus to be alive to the sorts of difficulties raised in *Ancient Medicine* (II.1*d*, **93–4**). But while Theophrastus goes on to deny that the theoretical heat of a substance has anything directly to with its immediate phenomenal properties (*Causes of Plants* 1. 21. 5), he maintains that it should be identified on the basis of a causal criterion: things theoretically hot *induce* heat; moreover, heat which can be perceptually identified (23 6).[35] This criterion provides the basis for

subsequent attempts (notably those of medical humoral theorists such as Galen) to supply non-arbitrary, empirically sensitive characterizations of their fundamental causal qualities.

On a different tack, Theophrastus claims (in *On Winds* 59) that explanations are particularly appropriate in cases where the phenomena themselves are intrinsically surprising; by contrast, what is reasonable will generally be accepted without demur. And he is not, in such cases, averse to stating a variety of distinct (and sometimes incompatible) *possible* causes for some phenomenon, if its actual cause is difficult to establish:[36] in this he anticipates the attitude of the Epicureans, in particular to the explanation of meteorological phenomena (VI.2*f*). Indeed, the recent recovery of the Arabic epitome of Theophrastus'*Meteorology* confirms the closeness of the links between his views and those of Epicurus,[37] and is in any case of independent interest in reconstructing his conception of the role of the divine in the explanation of the universe. Aristotle's God leads the world only by way of example (§2*g* above); he does not intervene directly into the management even of the celestial mechanism, much less that of the terrestrial sphere. Theophrastus too wished to restrict the extent to which appeal to the divine should function in explanation: but his restrictions do not exactly parallel Aristotle's.

Theophrastus explicitly denies that God has any role in the production of such traditional divine phenomena as earthquakes and thunderbolts; in this, of course, he is squarely in the tradition of naturalistic explanation that goes back to Anaximander, perhaps even Thales (I.2*b, g*). Aristotle too had offered severely mechanical explanations for such phenomena (although he did not shrink from appealing to an immanent teleology to explain regular meteorological phenomena: IV.3*d*; cf. §2*d* above). But Theophrastus sets himself both directly to explain the occurrence and distribution of such phenomena, and to show that these patterns are inconsistent with their explanation as expressions of divine displeasure:

205 when someone demands that we give the reason why thunderbolts are more frequent in the spring, we can answer: For the thunderbolts to arise, clouds, wind, and fire are required; during the winter clouds and winds exist, though not much fire because of excessive coldness; but during the summer there is much fire, whereas no clouds can be found. However, because spring is a

end p.191

moderate, temperate time, one can find in it enough clouds and winds. Moreover, when someone demands that we give the reason why thunderbolts are more frequent in high places, we may answer that it happens for two reasons: firstly because in high places there are many winds and clouds, and the thunderbolts come from them; but secondly because high places are close to the clouds, but low places are remote from them. Therefore thunderbolts reach the high places before they are dissipated; they do not reach the low places because they are dissipated before that. (Theophrastus, *Meteorology* [6.] 67–81, trans. after Daiber)

That passage exemplifies some of the methodological constraints upon explanation we have already considered, notably that the physical account of the phenomenon must be able to explain (in conjunction with other initial conditions) why it occurs as and when it does: the two conditions requisite for the production of thunder and lightning are present only in the spring.[38] Moreover, a physical explanation, entirely consistent with this naturalistic explanation, is readily available to account for the prevalence of lightning-blasts on high ground.

So far, so good: but Theophrastus has another, equally important, theological axe to grind. The perennial folk-conception that lightning represents divine displeasure was perhaps receiving philosophical backing at this time from the early Stoics;[39] at all events, the Stoics were happy to allow the possibility (perhaps even necessity) of divine portents (VII.2*d*). Theophrastus is concerned directly to combat such suggestions:

206 neither the thunderbolt, nor anything else that has been mentioned, has its origin in God. For it is not correct to say that God should be the cause of disorder in the world: on the contrary, he is the cause of its arrangement and order, and that is why we ascribe arrangements and order to God and the disorder of the world to the nature of the world. And, moreover, if thunderbolts originate in God, why do they occur mostly during spring or in high places, but not during winter or summer or in low places? Furthermore, why do thunderbolts fall on uninhabited mountains, on seas, on trees, and on irrational living things? God is not angry with those. It would be even more astonishing that thunderbolts can strike the best people and those who fear God, but not those who act unjustly and propagate evil. It is thus not right to say about hurricanes that they come from God. ([14.] 14–26, trans. after Daiber)

The hypothesis that thunderbolts exhibit supernatural wrath and punishment will not account for their distribution (if they did, they should be much better targeted), nor for the patent fact, so worrying to

end p.192

countless generations of believers who have perversely refused to accept Theophrastus' argument, that sinners' ways prosper. That argument recalls the central contention of *The Sacred Disease* (II.1*a*); and it is similarly persuasive. But note that Theophrastus does not reject the hypothesis of divine control over the Universe altogether: indeed the claim that God is ‘the cause of its arrangement and order’ perhaps suggests (although it does not entail) that Theophrastus envisioned a more active role for his God than had Aristotle.[40] At all events, order and structure are in some sense to be attributed to the divine (this *may* amount to no more than the familiar Aristotelian denial that pure, blind mechanism can produce order). Anything disorderly, on the other hand, must be attributed to the nature of things, presumably their material nature. Theophrastus thus injects a theological twist into the familiar Aristotelian notion that matter resists the imposition of form (IV.1*a*).

(b) Teleology and Metaphysics

On the other hand, Theophrastus perhaps goes further than Aristotle in restricting the role of teleological explanation. We have already seen how Aristotle refuses to allow that absolutely everything has a teleological explanation: some events are pure coincidences (IV.2*b*), while some animal structures are mere by-products, useless in themselves (IV.3*d*). Theophrastus is even more insistent that some common occurrences which exhibit no regularity (and hence cannot be explained purposively) are simply side-effects, occurring perhaps of necessity, but to no end (*Metaphysics* 9. $10^{a}26$–7); by contrast what is attributable to God must be ‘invariably uniform and similar’ (9. $10^{b}20$).[41] The *Metaphysics* is difficult to interpret, since it is avowedly aporetic in tone, raising problems and difficulties rather than resolving them;[42] yet we may still attempt to discern some of Theophrastus' own doctrines behind its problematic façade.

The problem dealt with by the *Metaphysics* is stated in the first paragraph:

207 how, and with what methods, must we delimit the study of first principles? For that concerned with nature is more varied, and, as some at least put it, more disorderly, covering every kind of change. That concerning first principles is determinate, and always about the same things. (Theophrastus, *Metaphysics* 1. 1)

end p.193

Theophrastus is concerned with a profound problem: how can the apparent variety and disorderliness of nature be reduced to the perfect certainties embodied by the basic explanatory principles? It is, he says ‘more reasonable’ to suppose that some connection exists between events, that there is more to the universe than mere chance (1. 2); and thus that there are indeed some prior principles. But what form can they take? If they are purely mathematical (the view of the Pythagoreans, and of Xenocrates, third head of the Academy: XI.1*b*), then it is hard to see how they can get any grip on the world of sensation (1. 3). Whatever the ruling principle is, it must be divine (1. 4), and a cause of movement, although itself unmoving, by virtue of arousing a desire to emulate it (1. 5).

So far, so Aristotelian: yet there are severe problems with this. The nature of the desire involved is unclear; and in any case the heavenly bodies rotate differently, and without end, so the object of their desire is never achieved (2. 7; cf. 10). Moreover, why should they not better imitate the best object by rest, rather than motion (cf. 5. 16)? In any case, desire presupposes soul, and hence the heavenly bodies would themselves be the causes of their motions (2. 8; cf. 8. 27–8, 11. 28; **210**). Finally, is motion essential to the heavens? If so, then the desire upon which it depends cannot be incidental to it (2. 11). These objections are not part of an isolated attack on the Aristotelian picture: similar, more devastating, criticisms can be mounted against the Platonists (3. 11–13). But they illustrate a general point:

208 so starting from this first principle (or these several principles), one might well demand that they should immediately generate what follows from them, and not to stop after advancing a certain distance (and the same goes for any other principles that anyone cares to postulate). (2. 11–3. 11)

Then again, should the principles be conceived as formal or merely material? In the latter case the problem of explaining the emergence of order from chaos re-emerges (4. 14–15); yet

209 on the other hand it is difficult to assign accounts to each individual thing by relating them to a final cause in all cases, both in animals and in plants and in the bubble itself;[43] unless it happens that the many and various forms of things in the region of the air and earth occur because of the ordering and changing of other things, the greatest example of which, according to some, is what has to do with the annual seasons, on which depend the generation of animals, plants, and fruits, with the sun as the generator, as it were. (4. 15; cf. 9. 30)

Explanation is a matter of discovering generalities (8. 20); and it is important to get clear about how many different sorts of knowledge (of

particulars as well as universals) there are (8. 22–5). Causal speculation begins from particulars discerned by the senses, but it cannot remain with them—yet it is difficult to see how the ascent can be made to the general principles (8. 25–6). The teleological doubts of **209** are expanded in the final chapter:

210 concerning the notion that everything is for the sake of something and nothing in vain . . . some things in particular do not seem to be so, but rather to be either coincidental or the result of necessity, both in the case of the heavenly bodies and in most of the terrestrial cases. For what is the purpose of the approaches and regressions of the sea, of the extension of land, of droughts and floods, and generally of things which change now this way now that, of destructions and generations, and plenty of other similar things? Moreover, even in animals, some things are next to useless, for example male breasts, and the female emission (unless this does contribute in some way), the growth of beards in some animals or in general of hair in certain places. (11. 28–9)

This passage is striking for a number of reasons. First, Theophrastus suggests, against Aristotelian orthodoxy, that even some heavenly phenomena are non-teleological in character (perhaps an indication that ultimately he would reject the unmoved mover in favour of an explanation simply in terms of the bodies' several natures).[44] Second, he picks on some regular occurrences which seem even so to exhibit no purpose: yet it is a feature of Aristotle's system that regularity implies purposiveness (IV.1*b*: the extent to which Aristotle's teleology is immune from such criticisms turns on the extent to which he is prepared to make use of the residue story: IV.3*a, e*). Finally if the 'female emission' is the menstrual flow,[45] Theophrastus implicitly at least calls into question Aristotle's account of generation (where it plays the crucial role of supplying the foetal matter: §2*a* above; IV.2*a*, 3*b*). Moreover, he goes on to instance those deer whose horns are a positive encumbrance, a case mentioned by Aristotle (cf. IV.3*e*).

Even so,[46] these are not cases which Aristotelian teleology is incapable of handling: on the contrary, as we have seen, it can do so rather well. On the other hand, Theophrastus' doubts about the PM, and his corresponding tendency to look for internal explanations for such things as the structure and motion of the heavens, reveal him as being less comfortable than his master with appeals to the hierarchical ordering of things within the overall structure of the universe (of course, there can still be hierarchies of function within particular organisms). And this is borne out by a final set of considerations: even if we want to say that 'nature in all things desires the best' (ibid. 9. 31), we should not do so without qualification.

end p.195

Theophrastus gestures towards Aristotle's explanation of the relative positions of windpipe and oesophagus, and of the corresponding necessity of the epiglottis (IV.3*e*); and at the very least suggests that such explanations seek to do too much.

In the end, from our standpoint, it matters little whether we may with certainty divine Theophrastus' own viewpoint in his *Metaphysics*: for he undeniably raises problems of fundamental importance to the business of explanation; and the subtlety of his admittedly brief discussion is evidence of a new sophistication in the handling of the notions of explanation and understanding. Theophrastus poses with unprecedented sharpness (and admirable succinctness) the basic questions which any moderately ambitious programme of explanatory science must seek to answer.

(c) Aristotelianism Refined: Strato of Lampsacus

On Theophrastus' death in 287 bc , the leadership of the Lyceum devolved upon Strato of Lampsacus, a post he held until his own demise some twenty years later. Strato was styled 'the naturalist', so Diogenes says, 'because above anyone else he devoted himself most assiduously to this study' (*Lives* 5. 58; cf. 64); but a passage of Cicero suggests another reason for the sobriquet:

211 you [i.e. the Stoics] say that nothing can be without God. But here Strato of Lampsacus attacks your flank, granting God immunity from any great labour. . . . he says he makes no use of God in constructing the world. He teaches that everything that exists, of whatever sort, has been produced by nature, although not in the way of him [i.e. Democritus] who says that these things were put together from rough and smooth, hooked and twisted atoms interspersed with void . . . but, having run through the various parts of the universe, he teaches that whatever is or will be is or has been made by natural motions and weights. (Cicero, *Academica* 2. 120–1 = Fr. 32 W;[47] cf. 1. 34)

According to Cicero, Strato rejects Democritus' favoured atomism (see VI.1*a–d*), while still maintaining that mechanistic explanations are sufficient in themselves to account for the universe and all of its contents. He strongly rejects the type of divine teleology, in which a Creator-God on Plato's model (III.4*a*) actually constructs the world, which was being elaborated at this period by the early Stoics (VII.3*a–b*), and whose principal opponents were the Epicurean inheritors of Democritus' atomist mantle (VI.2*d*).

As far as we can tell, Strato attempted to account for the world's workings by the operation of a purely mechanical nature, whose internal forces

account for generation, growth, and diminution (Cicero, *On the Nature of the Gods* 1. 35 = Fr. 33 W; cf. Fr. 34 W), a nature which may be called, although presumably only metaphorically (perhaps even ironically: cf. Fr. 37 W), divine, since 'it lacks entirely sensation and form' (cf. *The Sacred Disease*: II.1*a*, **65**). Plutarch stresses his originality:

212 further, of the other Peripatetics, the pre-eminent Strato disagreed with Aristotle on many counts, and held opposing views to Plato concerning the mind, soul, and generation; and in fine he says that the cosmos itself is not an animal, and that the natural is consequent upon the fortuitous; for he makes the spontaneous the principle, then derives each natural occurrence from it. (Plutarch, *Against Colotes* 1115b = Fr. 35 W)

Strato thus rejects Aristotle's immanentist teleology as well as Theophrastus' limited acceptance of a role for the divine, subscribing instead to the view (by this period paradigmatically associated with the Atomists) that the emergence of structure is, in the final analysis, a matter of mere chance (VI.1*d*; cf. I.4*c*). That is not, however, to say that the universe is irregular, or ungoverned by physical mechanisms: it is rather to insist that those mechanisms on their own can provide a satisfying and complete account of its structure. Strato thus places himself firmly on the mechanists' side of the great dispute between teleology and mechanism which is to characterize the remainder of the history of ancient (and subsequent) scientific explanation.

Precisely what mechanisms Strato subscribed to, however, are not easy to discern, although several reports credit him with making the qualities of Hot and Cold primary principles (Frs. 42–8, 89 W), following the lead of Aristotle and Theophrastus (§2*e*; §3*a*, **204** above). He parted company, however, with Aristotle (see §2*b* above) over the nature of lightness, holding, with the Epicureans, that all bodies have positive weight, lightness simply being a relative lack of it (Simplicius, *On 'On the Heavens'* 1. 8. 267. 29–268. 4, 269. 4–6 = Frs. 50, 52 W; cf. Aëtius 1. 12. 7 = Fr. 51 W): thus fire rises not out of any natural tendency to do so, but rather because, being relatively less dense than the other elements, it is squeezed out upwards (cf. VI.1*c*), a view which obviously involves a

rejection of Aristotle's account of natural place (as well as involving a gain in explanatory economy). He further concerned himself with *aitiologiai*, or causal explanations, for large-scale physical phenomena: thus he tried to account for the currents through the Bosporus and elsewhere on the hypothesis that different stretches of sea were at different levels (Strabo 1. 3. 4–5 = Frs. 91–3 W). Such aetiologies are characteristic of the later Peripatetic school, and, in a less scientific sense, of the Hellenistic temper in general (compare the third-century Alexandrian poet Callimachus' collection of mythological *Aitiai*, or causes for things).
end p.197

Thus a concern with the explanation of physical facts and occurrences came to be seen as characteristic of the Peripatetic tradition, as exemplified in the vast, sprawling collection of *Problems* falsely attributed to Aristotle. The *Problems* essay explanations, frequently in a cautious and undogmatic tone, for phenomena as disparate as sympathetic yawning, contagious diseases, and the concupiscence of the hirsute. It is reasonable to suppose that some of the collection reflects the tenor of the research programme in Strato's Lyceum (part of their cautiousness consists in suggesting, on occasion, several different possible explanations for something: here the Peripatetic tradition takes over the Theophrastean lead: §3*a* above).

Perhaps Strato's most influential views concerned void. He denied the existence of the Stoics' extramundane void (Aëtius 1. 18. 4 = Fr. 55 W; cf. Fr. 54 W), and also held, contrary to the Epicureans, and in good Aristotelian fashion, that space was a plenum (Simplicius, *On the 'Physics'* (*Corollary on Place*) 601. 22–4, 618. 20–5 = Frs. 59–60 W). Void was a hypothesis unnecessary to explain the fact of motion (cf. I.3*e*, 4*c*; VI.1*a*, 2*c*):

> **213** Strato's example offers an easier escape-route from these problems: whenever anyone places a pebble in a jar full of water, and inverts the jar while holding a stopper over its mouth, the pebble will be carried to the mouth of the jar while the water in turn takes the place of the pebble. (Simplicius, *On the 'Physics'* 4. 7. 659. 22–6 = Fr. 63 W)

The experiment is certainly striking and pertinent (compare his observational evidence for the view that everything moves faster the closer it approaches its natural place: 5. 8. 916. 14–28 = Fr. 73 W): but it will disconcert no atomist, who will simply maintain that the phenomena thus observed do not demonstrate either the non-existence or indeed the non-necessity of the void.

Such illustrations are designed to refute the view that void is a necessary condition of motion. Strato also argued against the necessity of the void to explain motion in a stronger sense, by way of the principle of *horror vacui*: the phenomenon of attraction, if indeed it exists, does not entail the existence of void; and in any case it is not clear whether there is really such a thing as attraction (4. 7. 663. 38 = Fr. 62 W). Moreover, Strato tried to account for magnetic attraction in a manner which made no reference to void, although the actual details of the explanation are obscure (4. 6. 652. 19–25 = Fr. 61 W; see further IX.1*c*; XI.4*d*). None the less, Strato held that bodies must in some sense contain void, since otherwise light and other powers could not permeate them (4. 9. 693. 11–18 = Fr. 65a W); Strato will thus not allow, as the Stoics were to, the possibility of total bodily interpenetration (see VII.1*b*).

Strato distances himself still further than Theophrastus from Aristotle's immanentist teleology—and he will, as we have seen, have nothing at all to do with any divine providence. He denied the necessity of the Aristotelian
end p.198

fifth celestial element, the ether (Aëtius 2. 11. 4 = Fr. 84 W: cf. §2*b* above); and equally his insistence on the sufficiency of internal forces does away with the necessity for postulating unmoved movers (Fr. 32 W; cf. §2*g* above). Moreover, he insisted (Simplicius, *On the 'Physics'* (*Corollary on Space*) 618. 20–5 = Fr. 60 W) that space was absolute, to be defined in terms of pure extension, and not merely the relative positioning of objects, against a suggestion of Theophrastus (639. 13–22 = Fr. 149 FHS&G: Theophrastean space derives its characteristics from the form of the objects it contains, rather than, in Aristotelian fashion, determining that form).[48] More controversially he may also have held time to be absolute as well. He certainly rejected Aristotle's definition of time as 'the number of change', on the grounds that rest takes place in time as well (Simplicius, *On the Physics* (*Corollary on Time*) 788. 36–790. 29 = Frs. 75–7 80–1 W); and he may have generalized this thought in order to show that time cannot be dependent upon change of any kind.[49]

Although styling himself a Peripatetic, and hostile to atomism as such, Strato has clearly moved a very long way from Aristotelian orthodoxy, in his (limited) acceptance of void, and his rejection of Aristotle's notions of place, space, void, and the Prime Mover, as well as his general teleology (to this list may be added Strato's un-Aristotelian view that the female contributes seed in generation: Galen, *On Semen* 4. 629 = Fr. 95 W). It is perhaps unsurprising in view of this that Aristotelianism went into decline in the period following his death. Strato's influence certainly continued to be felt, not least in the astronomical researches of his pupil Aristarchus (X.2*b*), and perhaps also in the physical views of the Alexandrian physician Erasistratus, whom the tradition at least connects with the Aristotelian school (IX.1*c*). But for the next two centuries the history of natural explanation is dominated by the clash between the atomism of the Epicureans and the continuous physics of their Stoic rivals, where once again battle is joined between the purveyors of pure mechanism and the partisans of a teleological providence. For that time and more, Aristotelianism languished; and it is not until the emergence of Alexander of Aphrodisias around the end of the second century ad that the philosophy of the Stagirite once more finds a worthy champion (X.4).

4. Conclusions

Aristotle sought to underpin his subtle discussion of explanation and its categories (IV.1–3) with a detailed and rigorous account of the structure of
end p.199

scientific knowledge: science should mirror reality in exhibiting what is causally derivative as being so derived from fundamental axioms about the structure of things (§§1*a–b*). This structure can be expressed by making the appropriate divisions in things, but cannot be discovered by any arbitrary 'method of division' (§1*c*). Rather the articulation of reality must be discovered by empirical investigation, and any theory is (or at any rate ought to be) subject to the canons of empirical adequacy (§2*a*). Consonant with this general methodological principle (but also with his deep commitment to the appropriateness of teleological explanation: IV.3), he develops a qualitative physics of motion based on the fundamental concept of the nisus of elements to their natural places (§§2*b–c*); he offers accounts of a wide variety of physical phenomena (§2*d*), and of chemical combination (§2*e*); and he sketches a cosmology designed to be at once both empirically adequate and teleological in cast without having it depend upon the conscious will of an intelligent designer (§§2*f–g*).
Aristotle's successor, Theophrastus, sought to fill in the gaps in the great Aristotelian programme of natural explanation (§3*a*); but he also developed and refines the methodology behind it, as well as insisting on the availability of non-supernatural explanations in the physical realm; and he is less sanguine than his master about the availability and appropriateness of teleological explanations (§3*b*), a tendency taken a stage further by his more severely mechanistic successor, Strato (§3*c*).
end p.200

VI The Atomists

R. J. Hankinson

1. Democritus

(a) The Background to Atomism

By the time of Plato and Aristotle there were, broadly speaking, two types of physical theory on the Greek market: the varieties of continuum theory adopted by Empedocles and Anaxagoras (I.4*b–c*) and the discontinuous world of the Atomists, whose earliest champions were Leucippus and Democritus (I.4*d*). Of Leucippus we know very little (he probably flourished *c*.440–430 bc); but the fragmentary tradition has served his younger contemporary Democritus better. Leucippus is credited with being the first to postulate the existence of atoms and the void (Cicero, *Academica* 2. 118) to circumvent the Eleatic denial of motion (Aristotle, *On Generation and Corruption* 1. 8. $325^{a}23–^{b}6$). Aristotle portrays him as accepting the basic Eleatic premisses
(1) what is is so absolutely;
(2) void is what is not;
(3) no part of what is is not.

He further agrees with the Eleatics that motion without a void is impossible, differing from them only in his attitude to (2): far from entailing that there can be no such thing, Leucippus treats it as *defining* void: void *is* absolute non-being (cf. *Metaphysics* 1. 4. 985^b4–10 = 555 KRS; *Physics* 1. 3. 187^a1–10 = 546 KRS). By contrast, the atoms that move in it are Parmenidean objects, eternal, unchanging, non-altering. Furthermore, Leucippus rejected the Zenonian argument that every object must be composed of an infinity of parts, presumably by rejecting the assumption that magnitudes must be infinitely divisible.[1] He started by assuming the senses to be (at least to a limited extent) veridical, and they assure us that motion exists. Hence we must supply a physical theory that explains how motion is possible, something which can only be done by the postulation of the void.

Democritus was equally concerned to save the phenomena:

end p.201

214 Diotimus says that according to him there were three criteria, the first, the apparent, being that of the apprehension of the non-evident; for, as Anaxagoras says (for which Democritus commends him): 'what appears is a glimpse of what is unseen' [59 B 21a DK]. (Sextus Empiricus, *Against the Professors* 7. 140: 68 A111 DK: 510 KRS)

But his attitude to sense-perception was a good deal more circumspect than that of his colleague:

215 sometimes abolishes the things which appear to the senses, and says that none of them appears in reality but only in opinion, the reality in things being the existence of atoms and the void:

by convention sweet, by convention bitter, by convention hot, by convention cold, by convention colour: in reality atoms and the void (68 B9, cf. 125 DK)

(Sextus Empiricus, *Against the Professors* 7. 135 = 549 KRS)

Several other fragments (68 B 6–11 DK = 549–50, 553 KRS) claim that we do not know how things are in reality, since the senses deal only in perceptual qualities which are merely 'conventional', and not representative of anything in nature.

Sometimes he even suggests that the whole enterprise is hopeless:

216 in reality we know nothing: for truth is in an abyss. (Diogenes Laertius, *Lives of the Philosophers* 9. 72: 68 B117 DK)

He was acutely aware of the epistemological problems posed for a theory which held that reality was fundamentally, indeed radically, different from the world as revealed by ordinary perception. The difficulty comes in reconciling empiricism (which Democritus espouses) with the view that reason ultimately shows the senses to be utterly unreliable, thus destroying the naïve realist picture of a direct and unproblematic correspondence between perception and the properties of things. Galen makes the point as follows:

217 everyone knows the greatest charge against any argument is that it conflicts with what is evident. For arguments cannot even start without self-evidence: for how can they be credible if they attack that from which they took their beginnings? Democritus too was aware of this; for when he had brought charges against the senses, saying . . . [68 B9 DK = **215** above], he has the senses reply to the intellect as follows:

wretched mind, do you take your evidence from us and then try to overthrow us? Our

217 everyone knows the greatest charge against any argument is that it conflicts with what is evident. For arguments cannot even start without self-evidence: for how can they be credible if they attack that from which they took their beginnings? Democritus too was aware of this; for when he had brought charges against the senses, saying . . . [68 B9 DK = **215** above], he has the senses reply to the intellect as follows:

overthrow is your downfall. (68 b 125 DK = 552 KRS)

So one should condemn the unreliability of an argument which is so bad that its most persuasive part conflicts with the evident propositions from which it took its start. (Galen, *On Medical Experience* 15, 114 Walzer)

Reason needs perceptual material to work on; but reason undermines those very perceptual reports, concluding from an analysis of that material

that the world must be radically different from the way it appears. It looks smooth, continuous, coloured, and solid; but reason shows that it is lumpy, discontinuous, without colour, and riddled with holes. What is more, if atomism is correct, what seems stable and motionless is in fact a continual buzz of atomic activity. Hence sense-perception contains the seeds of its own destruction: if sense reports are true, then they are false: and so they are false. But if they *are* false, then the superstructure reason builds upon them has no foundation, since it was based on the *assumption* that they were true in the first place.

Democritus' discussion takes the form of a dialogue between senses and the mind; and the mind's reply does not survive. In KRS (412–13) it is reconstructed thus: sense-perception does not give us veridical information about things, since it engenders atomism which refutes naïve realism. But perception still confirms atomism (in the sense that it is consistent with it, given Democritus' relativist perceptual theory), since atomism provides a causal account of why things appear as they do to the senses. Even so,

218 that in reality we do not understand how each thing is or is not has been made clear numerous times. (Sextus Empiricus, *Against the Professors* 7. 136 = 68 B10 DK = 550 KRS)

(b) Primary and Secondary Qualities: Reduction, Elimination, and Emergence

The Atomists are thus empiricists, and the goal of their theory is to account for the world as it presents itself to the senses. But they find that this takes them far from the ordinary world of common sense: the properties of the macroscopic world are to be explained on the basis of the micro-properties of the fundamental components of the universe, the atoms, and the void in which they move.

Atomism then involves some sort of ontological reduction. But the fundamental properties need not all be similar in type to the ones which they are invoked to explain; and in some cases (particularly those to do with perceptual properties) the apparent

macro-properties are simply emergent upon suitable arrangements and configurations of atomic qualities. Atoms really do possess solidity, resistivity, and weight as well (*On Generation and Corruption* 1. 8. 326^{a}9–10; Theophrastus, *On Senses* 61; Simplicius, *On 'On the Heavens'* 712 = 573–5 KRS);[2] and these properties are invoked to explain the creation of macroscopic objects, and their behaviours. By contrast, atoms do not have colour, taste, or temperature in
end p.203

and of themselves (**215** above; cf. Lucretius 2. 730–864). We do not know precisely how Democritus accounted for perceptual properties, except in the case of taste, where Theophrastus (*On Senses* 66 = 591 KRS) reports that bitter flavours are caused by 'small, smooth, rounded atoms' while saltiness is the result of larger, jagged atoms (Epicurus offered fuller explanations along the same lines: *To Herodotus* 51–4; cf. Lucretius 4. 615–72).

Thus whether or not something tastes sweet depends upon a set of facts about its structure that are not themselves phenomenal at all. This supplies the sense to his claim that things are merely by convention hot and cold, sweet and bitter, and so on (**215**). But there are in principle two distinct ways in which this might be cashed out.

First of all, phenomenal heat and cold may be 'conventional' for the same reasons that informed Protagorean relativism (II.2*a*): one and the same thing (the wind, for instance) can seem warm to one person, cold to another. That in turn is susceptible of a variety of interpretations: but it at least points towards an elaboration of a causal theory of perception in which both the states of the observer and the actual properties of the thing observed contribute to the particular perceptual property as it is observed by an individual. Thus, to say that honey is sweet by convention is not to assert that it is merely a matter of arbitrary fiat whether or not it is sweet, or even that its sweetness is a purely relative property. Rather phenomenal sweetness is the functional result of the combination of the dispositional properties of the substance, and the condition of the observer (in what was to become a standard sceptical trope, honey does not appear sweet to jaundice-sufferers: Sextus Empiricus, *Outlines of Pyrrhonism* 1. 101). Such a position is not proprietary to atomism; any moderately sophisticated causal-dispositional theory of perception will adopt it, regardless of ontology and metaphysics.

But there is a second sense in which such properties may be 'conventional': even honey's dispositional power of affecting the normally constituted as being sweet is not itself directly a property of the atoms in question. Rather honey is (dispositionally) sweet in virtue of the properties of its constituent atoms. And this claim too may be variously interpreted. The reductivist interpretation holds that sweetness is a derivative property: atoms fundamentally have only those properties that are the direct result of their being extended and solid (shape and size). For all that, the reductivist may allow that individual atoms are (in a sense) sweet: they have that property which, in conjunction with a suitable observer, induces perceptual sweetness; and nothing in this view commits them to asserting in any sense the unreality of perceptual sweetness.

On the other hand, the eliminativist holds in addition that the perceptual property is fundamentally unreal: all there really is exists at the atomic
end p.204

level: and whatever else atoms are, they aren't sweet. It seems probable that Democritus took this hard line: at all events, it squares well with the uncompromising, robust rejection of the reality of anything other than atoms (and their immediate properties) and the void (**215**), as well as making good sense of the scepticism (**216–18**). Alternatively, one might reject any form of reductionism, holding instead that a rather weaker relation obtains between individual atomic properties and perceptual affects. Thus the Epicureans were to contend that perceptual properties (or at least some of them) were merely emergent upon collections of suitably configured atoms, and hence in so far as the secondary qualities can be attributed (by way of a dispositional theory) to the atoms, it can be only to aggregations of them. Even construed dispositionally, colours and the like are not properties of individual atoms themselves (a similar account was no doubt given in the case of soul or mind); and herein perhaps lies the major difference between atomism in its Democritean and its Epicurean forms—but that is a story for later sections (§§2*b*, 3*b*).

(c) Atoms and the Void

At its most basic level, the Atomists' world consists of void, absolute non- being which none the less in some sense exists, and the indivisible particles or atoms which occupy it. The atoms are of different sizes and shapes; and the various properties of macroscopic stuffs are explained in terms of these shapes and their combinatorial properties. Thus mind and fire, being volatile, are composed of very small, spherical atoms (Aristotle, *On the Soul* 1. 2. 405^a8–13 = 585 KRS; cf. 591 KRS); and Democritus held that the basic properties of other stuffs could be similarly explained, although he apparently did not (unlike the Epicureans) actually offer any specific account of them (Aristotle, *On the Heavens* 3. 4. 303^a10–16). The atoms are not themselves the minimal units of size (except perhaps the smallest ones): atoms have parts (otherwise they could not be distinguished by shape),[3] even though these parts are not separable (or else their fundamental and Parmenidean status would be compromised). Thus Atomists need to explain why, even though they are not (notionally) partless, none the less the atoms cannot be broken into parts. Simplicius suggests an answer:
end p.205

219 they [i.e. Democritus, Leucippus, Epicurus] held the first principles to be infinite in number, which they considered to be uncuttable (*atomoi*), indivisible, and unaffectable owing to their compactness and their having no share of void; for they say that division comes to be because of the void in bodies. (Simplicius, *On 'On the Heavens'* 1. 7. 242. 18–21 = 67 A14 DK = 557 KRS)

It is not, then, simply their smallness which prevents them from being divided (as suggested elsewhere by Simplicius: n. 3 above), but rather the fact that they are completely compacted. Physical division occurs, according to **219**, along fault-lines of

void that exist even in the most apparently solid objects (compare the view of Strato: V.3*c*).
This account raises several pressing questions. First of all, **219** is at least compatible with there being atoms of any size whatsoever (and Democritus is said to have held that there was no theoretical limit to atomic size: 'it is possible for there to be an atom the size of the world': Aetius 1. 12. 6). Furthermore, it suggests that there is no differentiation in the basic stuff of atoms. This receives indirect confirmation from a text of Aristotle (*Metaphysics* 1. 4. 985^{b}4–19, = 555 KRS) which makes the distinctions between atomic types consist solely in differences of 'shape, arrangement, and position'.[4] But if the atoms are undifferentiated in regard to their stuff, they are absolutely distinct from one another in type: there is no transmutation of elements at the atomic level. In this regard the atomist account resembles that of Anaxagoras: apparent transmutations will in fact consist in the separating off of atoms of a certain type from the rest.
The atoms have weight (unless Aëtius is right: above, n. 2), presumably because it made little sense to treat weight as an emergent property (we may consider Democritean weight here as equivalent to our concept of mass). But atoms differ in weight solely as a function of differences in their sizes (573–4 KRS; § *b* above); and that also strongly suggests that they are undifferentiated in terms of their actual material constituents. Differences of aggregate weight for volume only emerge at the macroscopic level: 'the lighter is that which contains more void' (Theophrastus, *On the Senses* 61). Fire (considered as an aggregate of atoms) has weight—but there is far more void between the atoms of fire than of aggregates of heavy substances; hence fire is 'squeezed out by things that possess more, moves upwards, and consequently appears light' (Simplicius, *On 'On the*
end p.206

Heavens' 712). Aristotle treated both heaviness and lightness as positive properties of things (*On the Heavens* 4. 1–6); the Atomists make lightness simply the absence of weight (as did Strato: V.3*c*).
This leaves the important issue of how void is to be conceptualized for the Atomists, and how they justified the claim that it exists, even though it is 'what is not'. There are two principal possibilities. Either (*a*) void is the unoccupied space between atoms (the usual interpretation); or (*b*) it is absolute space (thus there will be void everywhere, in some places occupied, in others not). (*b*) has recently been championed by Sedley (1982*b*), and is in some ways attractive; however, the account of division in **219** seems to favour (*a*). Presumably it exists, simply in that we can refer to it, and locate it even; but it 'is not' just in that it is not *anything*: it is not a kind of substance.
It is also unclear how the Atomists argued for the view expressed at the end of **219**: why think that divisions must follow lines of void? I suggest that their only reason for so hypothesizing was their obsessive concern to rescue plurality and motion from the grip of Eleatic logic. If we allow bodies to be continuous, then the Zenonian paradoxes of plurality (I.3*e*) threaten to destroy the coherence of our most basic concepts. Atomism, for all its apparent counter-intuitiveness, at least promises a way out of the paradox of infinite divisibility.
Still, a Zenonian might reply that actual physical divisibility is not the issue—what matters rather is notional divisibility. If that is still possible *ad infinitum*, the paradox

resurfaces in a different form. That too can be countered with the hypothesis of quantized space—but whether or not the early Atomists were prepared to go that far is unclear.

(d) The Generation of Worlds: Chance and Necessity

So much for the basic properties of the fundamentals: the Atomists still need to show how the world and all its contents can be generated on the basis of them.

220 The worlds come to be as follows: many bodies of all shapes move 'by excision from the infinite' into a great void, and coming together there produce a single vortex in which, by collision with each other and revolving in every way they are separated apart, like to like. When they are no longer able to revolve uniformly because of their numbers, the fine ones move towards the void outside, as if sieved, while the others remain together, and interweave and run together with each other, making first a spherical system. (Diogenes Laertius, *Lives of the Philosophers* 9. 31 = 67 A1 DK = 563 KRS)

That is the beginning of the cosmogony attributed to Leucippus (as the process continues, a 'membrane' is formed around the outside, while more

end p.207

atoms settle into the centre; the membrane draws in more material from without, and parts of it coalesce and then ignite to form the heavenly bodies), and which was taken over by Democritus.

The details are highly obscure. First of all, it is unclear how the vortex is supposed to be formed. Elsewhere the impression is given that the Atomists try to generate absolutely everything out of the random rebounding and intertwining of the atoms: the atoms, possessed of solidity and weight, bounce off each other as a result of collisions, and so generate more collisions in an endless sequence (Alexander, *On the 'Metaphysics'* 1. 4. 36. 21 = 580 KRS; cf. 577–82 KRS). However, some atoms are hooked, and as such are able to intertwine with one another to form more or less stable compounds (Simplicius, *On 'On the Heavens'* 1. 10. 295. 11–20, 1. 7. 242. 21–6 = 583–4 KRS); which in turn can amalgamate into larger agglomerations. Democritus apparently seeks to generate the cosmos simply on the basis of this small set of physical and spatial properties: resistance, solidity, and certain physical configurations. But if that is the case, it becomes difficult to interpret the cosmogony of **220**, which apparently appeals to some more abstract principle of like-to-like attraction.[5]

This brings out a general difficulty for atomism. If Democritus really does wish to pare his ontological resources down to the bare minimum, and generate everything out of initially random interactions of free-floating atoms in an endless void, then he is faced with the powerful objections of Aristotle: how on earth can such exiguous material generate the complex and ordered structure of the world as we know it? On the other hand, if he is tempted to introduce abstract general forces, such as a like-to-like principle, then the clean, mechanistic lines of the theory are compromised.

Equally, do the vortices just arise as a determinate result of random atomic interaction (in which case Democritus needs to be able to show how such order can arise out of chaos)? Or is the vortex tendency itself primitive? Aristotle certainly took the Atomists to mean the former:

221 there are some who make chance responsible both for this universe and for all the worlds, since a vortex arose as a result of chance and the movement which by separation brought everything into its present formation. (Aristotle, *Physics* 2. 4. 196^a24–7 = 586 KRS)

And it seems on balance more likely that Democritus sought at least to generate these features from the minimum of conceptual resources. He
end p.208

appeals elsewhere to the illustration of what happens in sieves (n. 5): what he has in mind is the fact that, under certain conditions, the sum of a set of apparently completely random motions and interactions (of the pebbles in a sieve, or on the seashore) can result in the emergence of order. Atoms tend to agglomerate in types not because of some innate and irreducible tendency to seek out their own kind, but simply as the dynamic long-term result of a microscopically chaotic system.

These things happen by chance; but for Democritus they also happen as a result of necessity (Diogenes Laertius, *Lives* 9. 45). How can this be so? A number of solutions are possible. First of all, the formation of worlds might be chance in the sense of non-sequential: it exhibits no large-scale pattern. Yet it is the product of necessity in that the sequence of events at the micro level is determinist in character. Alternatively (but not incompatibly), the formation of a cosmos may be chance in the Aristotelian sense (IV.2*b*): it is to be analysed as a matter of the intersection of the appropriate causal chains, an intersection not to be explained or predicted from *within* any such chain, and one which is not susceptible of explanation in terms of any final cause (cf. n. 7 below). It is on the other hand necessary in that the unfolding of each chain is necessitated, and hence so are the intersections between them.

Again, the physical structure of atoms determines (i.e. necessitates) a certain range of outcomes, and precludes others (for instance, there can be no atomic interpenetration): and the formation of a cosmos is circumscribed by those physical possibilities. For all that, the generation of any particular world is random. Thus Aëtius remarks: 'Democritus means by it [i.e. necessity] the resistance, movement, and rebounding of matter' (1. 26. 2 = 567 KRS). On the other hand, Diogenes Laertius writes: 'everything happens in line with necessity, since the vortex, which he calls necessity, is the cause of the generation of everything' (*Lives* 9. 45 = 566 KRS), which in turn suggests that the role of the vortex is primary; and although it does not actually preclude the possibility that the action of the vortex may itself be explained reductively in terms of the motions of the atoms that constitute it, it does not suggest any such interpretation, and consequently offers no hint as to how one might be developed.

(e) *Ou Mallon* and The Principle of Sufficient Reason

Democritus was an infinitist. He held that our cosmos was only one of innumerably many; and the number of atoms within them was infinite (cf. **219** above). Indeed he probably thought that there were an infinite number of atomic types, and an infinite number of tokens of each type (**224** below). Why should he hold this view? The answer is summed up in a brief phrase which was later to become a slogan of sceptical epistemology: *ou mallon*,

end p.209

'no more'. The expression appears already in Plato: a beautiful girl is, by comparison with a goddess, not beautiful; hence she is in reality no more (*ou mallon*) beautiful than not (*Hippias Major* 287e–89d).

In the mouths of later Sceptics, *ou mallon* was used to avow sceptical indifference to an issue—some say *p*, some say not-*p*: hence *ou mallon p* rather than its opposite. Democritus, who appears to have been the first to use the phrase systematically, occasionally intends it to have a sceptical force (Aristotle, *Metaphysics*. 4. 5. $1009^{b}9$ ff. = 68 A 12 DK, on conflicting sense-reports). But elsewhere, Democritus'*ou mallon* is used to express the non-reality of certain perceptual properties in objects:

222 from the fact that honey appears sweet to some and bitter to others, Democritus . . . infers that it really is neither sweet nor bitter, and pronounces in consequence the formula *ou mallon*. (Sextus Empiricus, *Outlines of Pyrrhonism* 1. 13)

From our point of view its most striking application involves a version of the Principle of Sufficient Reason (PSR: see I.2*f*):

223 if what lies outside [our cosmos] is infinite, matter and worlds seem to be so as well. For why should they be here rather than there in the void? (Aristotle, *Physics* 3. 4. $203^{b}25$–7)

We do not know why Democritus took the void to be infinite; but the Epicureans were later to employ 'Archytas' argument'. If the universe is finite, of size *s*, it must either be (*a*) bounded or (*b*) unbounded. If (*a*), then bounded by something, and hence the totality of things is larger than *s*; but if (*b*), then what is to prevent one moving beyond its notional limit? After an application of (*a*), the same question may be asked of the enlarged universe containing the original of size *s* and its limiter; and that step is iterable indefinitely, or until (*b*) kicks in. Hence the universe is unlimited in size.[6] Lacking the technical apparatus required to distinguish the concepts of finitude and boundedness (which allows curved spaces to be both finite and unbounded, and hence to evade at least the obvious applications of Archytas' argument), that argument is of great intuitive power. Perhaps Democritus anticipated it. But perhaps he also argued that the notion of a limit to something intrinsically insubstantial made no sense—there is simply nothing there to be limited.

At all events, the hypothesis of the infinite void was taken to be basic; and then, by an application of PSR, **223** argues that, in default of any reason to the contrary, matter must be spread more or less evenly throughout it. Given that it is infinite in extent, matter too must be infinite in supply or else it would be spread infinitely thinly through the void, and

end p.210

hence could never coalesce to form worlds. The scope of PSR must obviously be limited here: the world quite clearly does exhibit differences of material density (which are supposedly explained by the vortices), and matter is not spread uniformly throughout it. Democritus presumably thought that there was a threshold minimum general density of atomic material necessary in any region of space in order for vortices to occur—and if matter were to spread entropically throughout the whole of space, those conditions could not be realized. But there was nothing to prevent matter from so spreading unless the overall supply was such as to ensure that no region of space could ever become too severely depopulated.

Similar considerations were urged (by Leucippus this time) in regard to the nature of atomic types:

224 he posited an infinite number of elements in perpetual motion, namely the atoms, and held that the number of their shapes was infinite, since there was no reason for them to be of one kind rather than any other. (Theophrastus, in Simplicius, *On the 'Physics'* 1. 2. 28. 8–10 = 560 KRS)

Unless you can give a good reason why the atomic shapes should be limited, they had better not be.

PSR has found defenders, in one form or another, throughout the history of philosophy (perhaps the most famous was Leibniz: *Monadology* §32); but it must be said that, if PSR is interpreted standardly as an a priori constraint on explanation (even if a synthetic one, as Kant held: *Prolegomena*, Preamble, §4), at the very least it compromises the pure empiricism of the Democritean picture. The Epicureans were, in general, even more explicitly concerned to portray themselves as empiricists: but they too will appeal to principles similar to PSR. PSR can, however, be interpreted as a second-order empirical hypothesis, one in principle open to question, if not outright falsification—and as such may function as a perfectly reasonable regulative principle on scientific methodology, provided its provisional status is respected. And there is some evidence, admittedly suggestive only, that later Epicureans may have attempted to construe PSR and its congeners in such a light. At any event, let us now turn to the Epicureans.

2. Epicureanism and Explanation

(a) The Principles of Epicurean Physics

Central to much of what has gone before is the notion that continuous processes demand explanation in terms of continuously immanent causes which keep them going: much of the peculiarity of Aristotle's dynamics is

occasioned by this demand. In general, the Greeks lacked the concept of inertia, the idea that a process will continue indefinitely unless something intervenes to alter it. Rather processes were the result of the exercise of powers, the failure of which would result in the cessation of the process.

It is worth noting here that their failure to come up with the notion of inertia is very probably a result of their empiricism: it just seems to be a brute fact that processes require causes, and that mechanisms run down—where they don't appear to, as in the case of the heavens, we need to posit an eternal cause of their eternal activity (V.2*g*). Here, the Greeks fail to carry through the process of conceptual idealization, of mentally stripping away the various factors which may be assumed, in the ordinary run of things in the ordinary world, to interfere with the smooth expression of a particular physical law. Indeed, perhaps Galileo's greatest claim to scientific genius (although it is only one among many), and one which sets him above the scientists of antiquity, is his ability to perform such conceptual idealizations. An object in free fall *would* accelerate uniformly at 32 ft./sec.^{2} *if* it were unimpeded by air resistance; and equally an object launched with a particular velocity would proceed for ever along its trajectory unless things intervened to prevent it from so doing.

At first sight, however, the Atomists from Leucippus and Democritus onwards appear exceptions to this general ancient rule. Their explanatory resources are pared down to a lean set of basic concepts, atoms and the void, and their respective fundamental properties. Everything that arises in the macroscopic world, whether object or process, is explained as the outcome of random[7] atomic encounters.

Epicurus (341–271 bc) took over the basic principles of atomist physics and epistemology from Democritus and his followers. But his outlook was distinctively Hellenistic, both in his pre-eminent concern with ethics (broadly construed as the study of how to live well), and in his fundamental interest in epistemology. Epicurus and his followers held that physics and metaphysics, indeed explanation in general, were of no use except in so far as they were conducive to the good life, which for Epicurus meant a life free from care and worry, the life of tranquillity, or *ataraxia*. Democritus too had preached the virtues of the untroubled life, of what he called ‘freedom from wonderment’ (*athambia*: 68 A 169 DK; cf. 68 B 215, 216 DK);[8] but it is Epicurus who draws the conclusion that learning and science, indeed the whole practice of explanation-giving, are valuable only in so far as they contribute to the individual's peace of mind. Epicurus
end p.212

develops his atomism with the express purpose of banishing the vain fears and superstitions that make human life a misery (and it is for this reason that Lucretius lauds him as an intellectual Hercules: 1. 62–79, 4. 1–54). Even the mechanistic, atomic cosmology has an avowedly ethical purpose.

But even if Epicurus' philosophy and science are subordinated to his ethics, none the less his views on explanation and cause (as preserved in the surviving letters *To Herodotus*, *To Pythocles*, and the fragments of his *magnum opus On Nature*, as well as by the first-century bc Roman philosopher-poet Lucretius) are of much independent interest.

First of all, Epicurus accepts the basic conservation principles that are the common property of post-Eleatic theorizing:

225 nothing comes to be from what is not, for in that case everything would come to be out of everything, with no need for seeds. Also, if that which disappears were destroyed into what is not, all things would have perished, for lack of that into which they dissolved. Moreover the totality of things was always such as it is now, and always will be, since there is nothing into which it changes, and since apart from the totality there is nothing which could pass into it and produce the change. (Epicurus, *Letter to Herodotus* 38–9 = 4a LS;[9] cf. Lucretius 1. 146–264)

Both creation from nothing and annihilation into nothing are conceptually impossible (see CP1, CP2: I.3*d*–*e*). The universe (of which our world is merely one among many: cf. Anaximander: I.2*f*), then, must always have existed. The fact that the universe is eternal gives grounds for the belief that none of its constituents are totally destructible, otherwise given an infinite past, they already would have been destroyed (this argument resembles that of Plato's *Phaedrus* 245e: III.2*a*, **131**(iv)).

The 'nothing from nothing' principle (cf. Lucretius 1. 146–214) may be construed either as CP1*a* ('nothing comes to be causelessly'), or as CP1*b* ('nothing comes to be except from pre-existing matter': I.3*d*). The Epicureans' concern in **225** is with CP1*b*; and they reject CP1*a*, at least in its full generality. Their opponents thought this inconsistent; Alexander of Aphrodisias (see further X.4) retails some anti-Epicurean Stoic arguments:

226 nothing in the world exists or happens causelessly, because nothing is independent of or insulated from everything that has happened before. For the world would be wrenched apart and divided, and no longer remain a unity for ever governed in accordance with a single ordering and arrangement, if an uncaused process were introduced. And an uncaused motion would be

introduced, if everything that exists or occurs were not to have some preceding causes from which it necessarily follows. For something to happen causelessly is, they [i.e. the Stoics] say, both similar to and as impossible as something coming to be out of what is not. (Alexander, *On Fate* 22. 192. 8–15 Bruns[10] = *SVF* 2. 945,[11] =55h LS)

That is, whatever counts in favour of CP1*b* equally supports CP1*a*. So why were the Epicureans moved at least to relax CP1*a*? And what form did that relaxation take?

Epicurus took over the basic ontology of atoms and the void from Democritus, albeit with one or two refinements. Epicurus, for instance, insisted that it was *not* possible for there to be an atom the size of the world, and that the variety of atomic types was finite, although very large (*To Herodotus* 42, 55). The reason for this is, no doubt, that given the doctrine of minimal parts (possibly already adopted by Democritus, but certainly refined by Epicurus: ibid. 58–9; Lucretius 1. 599–634, 2. 478–99),[12] there cannot be an infinite gradation of shape and size (the Atomists held that lines were composed of minimal parts—and hence there could be no such thing as a right-angled isosceles triangle: ps.-Aristotle, *On Indivisible Lines*).

More radically, he allowed that atomic motion was not invariably the result of their reboundings from other atoms. From time to time, an atom will swerve imperceptibly from its previous trajectory for no antecedent reason. And while Democritus' atoms move simply as a result of collision and momentum, Epicurus' are constantly falling downwards through infin- ite space (Lucretius 2. 184–215). Epicurus' reasons for this innovation are not clear; but he may have felt it necessary to inject some continuous dynamism into the physical system on the basis of the atoms' weight, construed as a tendency to move in a particular determinate direction (which we may call 'downwards'). Otherwise, in spite of the constant and eternal atomic interaction, the universe would simply run out of steam (cf. **225**). If this is right, Democritus came closer to postulating an inertial physics: but Epicurus was more in touch with the general Greek desire to account for the continuity of processes (cf. § *d* below).

(b) Perception and Emergence

Epicurus agrees with Democritus in restricting the fundamental properties of atoms to a small set:
end p.214

227 furthermore we must hold that the atoms possess none of the properties of phenomenal objects except shape, weight, and size, and whatever necessarily goes along with shape. (Epicurus, *To Herodotus* 54)

Yet his attitude to the status of the phenomenal properties possessed by macroscopic objects seems different. I argued above (§1*b*) that Democritus seeks to *eliminate* such properties from his ontology altogether. Epicurus, on the other hand, argues for, and attempts to account for, the *emergence* of phenomenal properties such as colours, odours, and temperature.

The atoms themselves cannot possess them, since they are variable, whereas the atoms by hypothesis are not (Frs. 29, 30, 288, 289 U;[13] Lucretius 2. 730–864); but they are not thereby rendered merely 'conventional' or unreal. Secondary qualities are rather the perceptual result of the various configurations of atoms, which can of course vary (*To Herodotus* 54–5; Lucretius 2. 730–990). Lucretius' procedure here is interesting. First of all he establishes that the idea of a colourless body is not incoherent (contrary to a widely held ancient view: cf. Plato, *Meno* 75b–c): 2. 730–48. Then he argues that atoms cannot be coloured, because things may change colour, yet atoms need to be unchangeable in order to underwrite the continuity of the world (2. 749–56). Next he contends that colours *can* be explained as being emergent upon arrangements and configurations of atoms (2. 757–67); finally he refutes the idea that the colour of macroscopic objects may change because atoms of differing colours become uppermost within them (2. 768–75), and rebuts the notion that the overall colour of an object may be the functional sum of the various colours of its atomic constituents (2. 776–87). Thus Lucretius attempts first to establish the possibility of the Atomists' position, and then to show how it is explanatorily preferable to any of the alternatives.

The mechanisms of perception are explained (again expanding upon a Democritean model) as involving different types of contact; distance senses work by direct contact not with the objects of perception themselves, but with atomic effluences from them. The Epicureans picture objects as giving off a continuous stream of particles in all directions (although they prefer to conceptualize them as laminae, solid films of integrated particles,

or simulacra: *To Herodotus* 46–50; Lucretius 4. 26–128), streams which by impinging on the various sensoria convey images to them, in a very real, indeed naïve, sense. We do not see the images themselves (ibid. 4. 257–8): we see by means of them (if they were to be visible, the images themselves would have to generate images). This is the
end p.215

basis of the notorious Epicurean doctrine that all perceptions are true (*To Herodotus* 50–2).[14]

Perception consists in a direct causal interaction between sense-organ and material, and so the actual act of perception itself cannot be false. However, perceptions can falsely represent their purported originals (images can become confused and damaged in flight, and they may entangle with other images, forming the thin composite images of dreams and the imagination: ibid. 51); and the mind can decide erroneously to accept them as genuine. Thus ‘all perceptions are true’ involves more than simply the truism that all perceptions are caused, without amounting to the hopelessly optimistic claim that we can never be perceptually misled.

(c) The Necessity of Atomism

Whatever its shortcomings, the Epicurean programme is of undeniable interest and importance. In the first place, it is extraordinarily bold and all- embracing; and the resourcefulness with which Lucretius, our most complete source, attempts to carry the project through attests to its importance for the Epicureans. Even so, the problems it faces are considerable. The ‘seeds’, of **225**, are pivotal. The principal challenge facing Epicurean, no less than Democritean, physics and cosmology is that of being able to generate a plausible account of the complexity and regularity of the cosmos on the basis of such a limited set of explanatory concepts. One of the standard ancient arguments for postulating causes in the first place was the very uniformity of nature's processes:

> **228** that cause exists is plausible; for how could there come to be increase, decrease, generation, destruction, in general change, each of the physical and mental outcomes, the ordering of the whole universe and everything else except by reason of some cause? . . . Moreover, if cause were non-existent, everything would have been produced by everything at random; thus horses might have been born of flies, and elephants of ants. (Sextus Empiricus, *Outlines of Pyrrhonism* 3. 17–18; cf. Lucretius 1. 159–73 = 4b LS)

The Epicureans attempt to solve the regularity problem by positing the existence of ‘seeds’, molecules of atoms organized in such a way as to act as templates for the accretion of similar aggregates (compare Anaxagoras' physics: I.4*b*). Some such mechanism is clearly required; but they do not, and perhaps cannot, provide any detailed account of how this might work.

In effect, the Epicureans try to show that it is only on the atomic hypothesis that the world of appearances can be accounted for. Motion is impossible without void (*To Herodotus* 40; Lucretius 1. 426–8); and infinite
end p.216

divisibility courts Eleatic paradox. Moreover, Epicurus suggests an argument congruent with **225**: if everything could be ground down into indefin- itely small parts, then in an eternity of time it *would* be so ground down; but nothing could then be rebuilt out of such a homogenous paste. Only by postulating indestructible atoms that will for ever potentially be able to intertwine and build up more complex bodies can we account for the continued existence of an organized world: and hence the seeds hypothesis must be true, even if its detailed structure remains a mystery.

That is a paradigmatic Epicurean argument; and evidently it stands or falls with the claim that the phenomena entail their preferred explanation:

229 everything is bodies and void. That there are bodies is attested in every way by perception itself, in accordance with which reason must judge what is obscure; and if there were no void, there would be no way in which bodies could move, as they are evidently seen to move. Besides these things, there is nothing which can be conceived either directly or on the basis of analogy. (Epicurus, *To Herodotus* 40)

The argument that nothing could move without void is rehearsed in Lucretius (1. 370–83). For a body to move at all, there must be vacant space in front of it or else its motion will be obstructed. The problem of motion in a plenum had already exercised the Presocratics, and Empedocles had essayed a solution to it, namely that of 'reciprocal circular movement' or *antiperistasis* (I.4*c*; cf. V.2*b*). The theory holds that movement in a (closed) plenum is possible provided that material ultimately describes a circle, matter slipping in behind the moving object, and in its turn having its former place occupied by more matter, and so on. Indeed, the process needs to be ultimately circular only if space is closed: otherwise nothing prevents an object simply pushing stuff in front of it, like guests in Hilbert's hotel.[15] And the atomist world *is* infinite; hence, even if they can show that *antiperistasis* is incoherent, they cannot infer on that basis that there must be void (although motion in an atomic plenum would be possible only for atomic shapes of a restricted type—e.g. cuboid—and then only in a few directions).

But in any case, what is wrong with *antiperistasis* theory? It fails to save the appearances consistent with a plenum only on the prior assumption that matter is atomic in structure. The proof of the existence of indivisible

particles of stuff is basic to the establishment of the whole theory. If that can be shown to be less than secure, then the necessity of the void falls with it. That is, things are less obvious than Epicurean sanguinity allows;[16] and there is no inference of the security they require from the way things seem to be phenomenally to the way they must be in reality (cf. VIII.2, **286–7**).[17] In other contexts, the Epicureans were happy to rest with merely probable, even disjunctive, accounts of how things might be. But in the case of the atomic theory, the Epicureans see their science as a matter of inference not merely to the best, but to the only, explanation (cf. Lucretius 2. 991–1022).

(d) The Shortcomings of Epicurean Reductionism

The Atomists are often commended for prefiguring modern styles of scientific explanation, and for seeing the possibility of explaining macro- properties as emergent upon typologically distinct micro-structures. But the Epicureans were unable to free themselves from the tendency to conceptualize the micro-properties as being similar to their macroscopic counterparts: ‘things that seem hard and stiff must be made up of deeply indented and hooked atoms’ (Lucretius 2. 444–5); while fluids are composed of smooth, rounded atoms. On the other hand, they are ready to allow that colour is purely emergent (2. 737–841), as are heat and cold (2. 842–64); and no one in their right mind would assume that there are laughing atoms (1. 915–20).

Atoms possess a natural, internal tendency to move ‘downwards’ (Epicurus, *To Herodotus* 61; Lucretius 2. 83–5, 216–18), not towards the centre of the cosmos (since there is none in an infinite universe), but along parallel vectors, a tendency which is invariant from atom to atom (ibid. 2. 225–42; the natural speeds of atoms in pure void are the same: *To Herodotus* 61; atomic speed is not, then, a function of weight, as Aristotle had it: *On the Heavens* 4. 1–6). The properties of solidity and shape explain their reboundings and intertwinings:

230 the atoms move continuously and for ever, some separating a great distance from each other, others keeping up their vibration on the spot whenever they happen to get trapped by their interlinking, or imprisoned atoms which link up. For the nature of the void brings this about by separating each atom off by itself, since it is unable to lend them any support; and their own solidity causes

end p.218

them as a result of their knocking together to vibrate back, to whatever distance their interlinking allows them to recoil from the knock. There is no beginning to this because atoms and the void are eternal. (*To Herodotus* 43–4 = 11a LS; cf. Lucretius 2. 80–124)

The last sentence hints that the obvious question of how the cosmic buzz of atoms got going in the first place is misguided, since there is no first place in a temporally infinite universe. However, later testimony uniformly ascribes to Epicurus a doctrine which is designed among other things to account for the origin of the collisions. Lucretius writes:

231 when bodies are being borne down by their own weight straight through the void, at quite uncertain times they veer a little from their course, just enough to be called a change of motion. If they did not have this swerve, everything would be falling downwards like raindrops through the depths of the void, and collisions and impacts among the primary bodies would never have arisen, with the result that nature would never have created anything. (Lucretius 2. 218–24 = 11h LS)

The notorious swerve violates CP1*a*, a fact which provoked some ancient derision (see §3*a* below). The fact that Democritus apparently saw no need of it, and its absence from the epitome of physical doctrines in *To Herodotus*, perhaps suggest that it is a late addition to the Epicurean physical armoury; and it probably owes more to Epicurus' determination to rescue human freedom from the clutches of universal determinism (§§3*a*–*b* below) than to any perceived lacuna in the cosmological story. The swerve was condemned for being *ad hoc* and explanatorily useless, as well as for violating CP1*a*. But to a modern sensibility less inclined to treat CP1*a* as being true a priori, Epicurus' innovation, whatever its shortcomings, represents a bold break with a suspect conceptual orthodoxy.

But whatever the nature and function of the swerve, Epicurean causal explanation is severely materialist. The ceaseless atomic bump and grind in the void explains both macroscopic events (which are ontologically parasitic upon bodies: Lucretius 1. 464–81) and macroscopic properties. All causing is, properly speaking, corporeal; void is not, then, a cause, but rather a prerequisite or necessary condition of causal interaction, since causation occurs by material contact, and hence requires movement, which is impossible without void.

It is tempting to view the Epicurean world of atomic interaction as involving something like an inertial physics: give an atom a shove in a particular direction, and it will continue to travel that way until it meets an obstacle, whereafter its trajectory will be determinable as the functional sum of the momentums and directions of the particles involved. But tempting or not, it is important to realize for the Epicureans the engine for action

is, no less than for Aristotle or the Stoics, a set of internal properties, or forces, in constant activity. This is why Epicurus introduces the constant fall. It is the weight of the atoms that makes them fall *and keeps them falling*.

Moreover, Lucretius does not conceptualize the physics of atomic collision in terms of mass and momentum, and the commensurate transfer of kinetic energy from one object to another. Rather when one object strikes another, it actualizes the latter's internal potentiality for motion. This is at least suggested by two passages. Lucretius explains the enormous velocity with which the simulacra peel off from the surface of objects as being the result of ‘a very slight initial impetus far behind them which launched them and propels them’ (ibid. 4. 193–4): here there seems to be no obvious conservation of energy between the cause and the effect. Elsewhere, Lucretius describes how the motions of the tiny, volatile atoms that make up the mind can amplify to produce animal movement, saying that this phenomenon is less surprising if we remember that tiny particles of wind may drive a ship (4. 886–91, 898–906). If this is right, then Epicurus, contrary to the usual exegesis, is conceptually much closer to the general Greek norm. Atoms move essentially under their own power, less like the billiard-balls of Humean analogy than runners in a relay-race, primed to take off at a touch.

(e) Teleology and Mechanism

But if the Epicureans' dynamics turns out after all to be less distinct from that of their opponents than might appear at first sight, they part company decisively with the great majority of ancient theorists in their rejection of teleological explanation in nature. No purpose animates the world or explains its generation; a cosmos is just one of those things that happens from time to time. Neither a Stoic–Platonic benevolent divine organizer, nor an Aristotelian internal drive towards actualization, are necessary to account for it:

232 one mistake . . . is that of supposing that . . . the eyes have been created in order that we might see; that it is in order that we might take lengthy strides that the knees and hips can be flexed above their base of feet . . . and hands supplied on either side as servants in order that we could perform whatever acts were needed for living. All other explanations of this type which they offer are back to front, products of distorted reasoning. For nothing has been engendered in our body in order that we might be able to use it. It is the fact of its being engendered that creates its use. (Lucretius 4. 823–35 = 13e LS)

In the dispute over the direction of explanation, the Epicureans take Anaxagoras' side against Aristotle: structure determines function, not the
end p.220

other way around. Lucretius explicitly states (4. 836–57) that all bodily organs (unlike artefacts, which are created to some purpose) exist prior to their functions.

The world is neither the product of a divine craftsman, nor was it created with human beings in mind (5. 156–69). The gods of the Epicureans enjoy a life of untroubled bliss in the interstices between the worlds; they have no interest in intervening in ours (*To Herodotus* 76–8). This is in the sharpest contrast with Stoic providentialist theology, which held that God permeated the entire world in the form of *pneuma*, the fiery air responsible for both intelligence and the cohesion of things (*SVF* 2. 526, 1027, 1077, etc.: VII.2*b*, 3*a*), and that everything in the world served some, frequently providential, purpose (VII.3*a*). The Stoics, like Aristotle before them, emphasized the enormous improbability of the generation of an atomist cosmos:

233 does it not deserve amazement on my part that there should be anyone who can persuade himself that certain solid and indivisible bodies travel through the force of their own weight and that by an accidental combination of those bodies a world of the utmost splendour and beauty is created? I do not see why a person who supposes this can happen does not also believe it possible that if countless exemplars of the twenty-one letters . . . were thrown into a container and shaken out onto the ground, they might form a readable copy of the *Annals* of Ennius. I'm not sure that luck could manage this even to the extent of a single line. (Cicero, *On the Nature of the Gods* 2. 93 = 54m LS)

The Epicureans countered by stressing the infinity of space and time, and the inevitability that in the long run all conceivable atomic arrangements will be realized. But even so, the Epicureans lacked the sophisticated explanatory tools necessary to show how a structured, regular, and persistent world and its denizens could emerge stably from the basic atomic chaos. And even now the argument from design proves itself remarkably philosophically durable.

(f) Multiple Explanations

Yet the Epicureans did not believe either that they could or that they needed to provide secure and definitive explanations for every phenomenon. After all, their pre-eminent concern was with *ataraxia*: the only importance of physics is to serve that ethical end by showing that the endemic fear and anxiety that plagues human life is groundless (§ *a* above). They cared little about knowledge for its own sake:

234 we hold that to arrive at accurate knowledge of the cause of the most important things should be considered the business of natural science, and that happiness depends on this. (Epicurus, *To Herodotus* 78)

end p.221

It did not matter if they could not produce incontrovertibly secure explanations for certain phenomena, such as lightning and earthquakes. Atomism was sufficiently well grounded for it to be certain that the ultimate truth about such things would be atomic; but precisely what that truth was of little import. All that was required to eliminate irrational human fears about them was to show that they were not the product of some malign supernatural influence:

235 there are also a number of things of which it is not enough to name one cause, but rather many causes, one of which will however be the actual one—just as if you were to see at a distance the dead body of a man, it would be appropriate to list all the causes of death, so as to include the specific cause of his death. For you would not be able to establish that he dies by the sword, from cold, from disease, or by poison; yet we know that something of the sort must have happened to him. And similarly in many other matters we are in a position to say the same. (Lucretius 6. 703–11 = 18e LS)

Epicurus, while maintaining that as regards the basic cosmological facts 'we must recognize . . . no plurality of causes or contingency' (*To Herodotus* 78), still holds that

236 when we come to subjects for special inquiry, there is nothing in the knowledge of risings and settings and solstices and eclipses and things of this kind that contributes to happiness. . . . hence if we discover more than one cause that may account for them . . . we need not think our account falls short in accuracy, so far as is necessary to keep us tranquil and content. (ibid. 79–80)

Epicurean 'explanations' of such phenomena are disjunctive in form: x occurs because either E_1 or E_2 or. . . E_n . At most one of the E_i 's can be the true explanation (cf. Lucretius 6. 703–4); but if the disjunction is suffi-ciently all-embracing, one of them will be: and that is all that is required. Epicurus' letter *To Pythocles* concerns 'celestial phenomena'. Having said once more that physical explanation serves to alleviate fear, Epicurus writes:

237 we do not seek to wrest by force what is impossible, nor to understand all matters equally well, nor make our treatment always as clear as when we deal with human life or explain the general principles of physics, for example that everything consists of bodies and intangible nature, or that the ultimate elements of things are indivisible, or any other proposition which admits only one explanation of the phenomena to be possible. But this is not the case with *ta meteōra*: these admit a multiplicity of causes of their coming to be and explanations of their nature consonant with perception. For we must not do science by way of empty assumptions and arbitrary fiat, but as the phenomena demand. (Epicurus, *To Pythocles* 85–7)

end p.222

The last sentence recalls the opening of *Ancient Medicine* (II.1*d*, **86–7**). Explanation must start from the appearances. Sometimes only one possible explanation can be given consistent with the empirical evidence; then that evidence entails the explanation. But elsewhere things may not be so clear-cut, and a variety of incompatible explanations could save the appearances: here multiple explanation is in order.[18]

3. Freedom and the Explanation of Action

(a) Bivalence and Determinism

Aristotle held that if the future was not contingent, then deliberation would be a waste of time (IV.4*b*), and human beings could not be free. But this posed a problem for the status of statements about the future: how can they now be either true or false, if the future is itself uncertain? On the other hand, how can neither of a pair of contradictory propositions about the future be true?

The Epicureans attacked this dilemma by holding that the principle of bivalence (that a proposition must take one—and of course only one—of the two truth-values true or false) failed for contingent statements about the future, which were neither true nor false prior to their being actualized (and *a fortiori* neither necessarily true nor impossible: Cicero, *On Fate* 21). Cicero further says that the swerve was explicitly introduced in order to underwrite physically the rejection of bivalence which was supposed to be essential for the maintenance of human freedom:

238 Epicurus introduced this theory [i.e. of the swerve] because of his fear lest, if the atom was always carried along by its natural and necessary gravity, no freedom will be left for us, since the mind will move under compulsion from the atoms. (Cicero, *On Fate* 23)

Thus the Epicureans attempted to rescue freedom and contingency. But they do so at some cost. First, the introduction of the minimal swerve blurred the clean lines of their physics. And secondly, the denial of bivalence is at the very least counter-intuitive. Even if Caesar could have escaped assassination, the claim ‘Caesar will not be murdered’, whenever

end p.223

uttered prior to the Ides of March, 44 bc , is actually false. If Teiresias or Elijah had said centuries earlier ‘Caesar will be murdered on the Ides of March’, surely they would have been telling the truth. The Stoics, impressed by the force of such considerations, accepted bivalence, and thought that they were thereby committed to some form of determinism (VII.1*a*, *f*, 2–3*passim*). Both Stoic and Epicurean, then, appear to hold both that

(1) if (*a*) truths are eternally fixed, then (*b*) if a proposition *p* is true at any time, *p* is true at all times;

and

(2) if (*a*) *p* is true at all times, then (*b*) *p* is necessary.

The Stoics argue from the truth of (1*a*) to the truth of (2*b*); the Epicureans contrapositively from the falsity of (2*b*) to the falsity of (1*a*). Neither position seems entirely comfortable.

The great Academic sceptic Carneades (*c*.219–*c*.129 bc) saw that the argument from (1) to (2) need not be compelling. We may distinguish the following two theses:

(T1) each event is the wholly determinate product of prior causes

and

(T2) if an event is going to happen, it is already true that it will happen.[19]

The Stoics suppose that if some proposition about the future is now true, there must now be some truth-maker for it in the world, and so it will already be unavoidable. But T2 does not entail T1 (although the converse entailment holds). The principle of bivalence may perfectly well hold for all future contingents without that having any implications for the current causal state of the world:

239 the truth of propositions like ‘Cato will come into the Senate’ is brought about by contingent causes, not by causes bound up in nature and the world. And yet that something will come about, when true, is as immutable as the truth that something has come about. (ibid. 28 = 70g LS)

If it is now true that I shall die of a heart-attack, then nothing I now do will change that truth (since truths are timeless). But this is not to say that my

end p.224

death by infarction is ineluctably fated (and hence that I might as well abandon myself to a binge of cigarettes, booze, and lard), since it will only now be true that I shall so perish if certain factors, some of which may well be within my control, concur to produce the coronary failure. If I succeed in preventing it from happening, then it will turn out that it never had been true that I was going to die of cardiac arrest.

Carneades argued as follows:

240(*a*) if all things come about through antecedent causes, all things come about through the interconnection in a natural chain. (*b*) If that is so, all things are the product of necessity. (*c*) If that is true, nothing is in our power. (*d*) But there is something in our power. (*e*) But if all things come about through fate, all things come about through antecedent causes. (*f*) Therefore it is not the case that whatever occurs does so through fate. (31 = 70g LS)

His main target was the Stoics (VII.2*d*); but the argument poses a challenge to any physical system that emphasizes the necessity of the causal order. It was precisely to counter arguments of this sort that Epicurus felt it necessary to introduce the atomic swerve; otherwise, he thought, the human freedom expressed by (*d*) (and presumably supported by ordinary experience) is fatally compromised.

However,

241Carneades . . . showed that the Epicureans could defend their case without this fictitious swerve. For since they taught that a certain voluntary motion of the mind was possible, a defence of that doctrine was preferable to introducing the swerve, especially as they could not discover its cause. And by defending it they could easily stand up to Chrysippus. For by conceding that there is no motion without a cause, they would not be conceding that all events were the results of antecedent causes. For our volition has no external antecedent causes. (ibid. 23 = 20e LS)

Carneades goes on to say that to act without cause (or freely, as we might say: cf. Locke, *Essay* ii . xxi. 7–27) means only to act without external compulsion. The causes of such actions are volitions, which are not, however, themselves caused. Thus the swerve is redundant for the Epicureans; and they need not accept uncaused events in order to reject determinism in regard to human action. They need not, then, 'incur the scorn of the natural philosophers' (Cicero. *On Fate* 25) by supposing that some events lack causes. But this appears problematic: if the volitions themselves (whatever they may be) have no external causes, then it seems that either they have no causes at all (in which case they cannot be events), or their internal causes must ultimately abut against something which is itself uncaused, and hence not an event. Perhaps in Spinozist fashion there are supposed to be 'pure acts of the will' (again, whatever they may be) which are (in some sense)

end p.225

self-caused. Views of this sort, although perennially popular in some circles, are notoriously obscure, and unconvincing to anyone impressed by Hume's argument that cause and effect must be distinct items. Moreover, the notion of a causally insulated, autonomous will is not easy to reconcile with the patent fact (accepted by Chrysippus and the Stoics) that external influences are necessary conditions of our having volitions at all. The following argument suggests itself: if autonomy involves having control over one's actions, then the will of the autonomous agent must be sufficient for those actions. But that appears to conflict with the admission that externals are necessary conditions for the action (except on the *outré* hypothesis that the volition is itself a sufficient condition for the existence of the external conditions). On the other hand, we may distinguish between what occasions the particular occurrence of the volition, the character of the volition itself, and its causal relations with subsequent action. The occurrence of a particular volition at a particular time may be conditional, in that only if certain external circumstances are realized will I will a particular course of action (only if I see the zabaglione will I form the volition to consume it). Even so, it may still be the case that when they are realized my volitions are sufficient for the outcome (nothing else is required over and above my volition now actualized in order to make me pursue the pudding); and furthermore, the impact of the external, occasioning cause (the perception of the custard) has nothing whatever to do with the formation of the disposition to consume such things which informs the volition itself. This, then, would be the sense in which ‘our volition has no external antecedent causes’; compatibly with this is it is also implied that external events are not the *complete* (fully sufficient) causes of our actions. But if this is right, it is hard to see why any causal indeterminacy should be necessary in the first place (which is exactly Carneades' point in suggesting an alternative, compatibilist account of how free action was possible on the basis of ‘voluntary motions of the mind’: **241**). And in any case, the Epicureans are left with the major task of showing how a random atomic event can make us free: randomness seems, if anything, even more inimical to freedom than determinism.

(b) Freedom and Causal Necessitation

But even if determinism cannot be directly inferred from bivalence, the anti-determinist still has much work to do. The principle that every event must have a cause is an appealing one: and if it is interpreted strongly (each event has a uniquely determining cause), determinism appears unavoidable, and causation will be assimilated to necessitation. Some modern philosophers (most influentially Anscombe 1975) have tried to deny the

link between the two; and some commentators ascribe such a denial to the ancients (notably Sorabji 1980*a*, *b*). Certainly Aristotle did not think his causes to be invariably necessitating: if *A*s are only for the most part causally correlated with *B*s, it may still be the case that this *A* causes this *B*, although it does not (on its own) necessitate it. But that still leaves open the question of whether events for Aristotle (considered as individuals, and not under some particular explanatory description) are uniquely determined by antecedent circumstances; and here the evidence is far more equivocal (IV.2*b*).

So if you think that universal causation entails a determinism incompatible with human freedom, but also take human freedom to be a given of ordinary experience, you will be forced to deny universal causation, as Epicurus and his followers apparently did. It is indeed fairly clear that Epicurus did think that determinism (of some sort at any rate) would (in some sense at least) be destructive of freedom (or at any event of the legitimate ascribability of responsibility to agents): but the questions still remain: *what* sort of determinism is ruled out, and *why* is it supposed to be so damaging? Carneades, as **241** shows, thought that the Epicureans solved nothing by introducing random swerves. Freedom cannot simply amount to arbitrariness. Free individuals are not merely loose cannons, firing uncontrollably.[20]

However, perhaps Epicurus did not take the swerve to be *constitutive of* human freedom. David Sedley has argued (Sedley 1988) that the swerve makes room for acts of volition, and need not necessarily be uncaused (although it is uncaused *by simple atomic interaction*). The fact that atoms can be deflected allows volitions to get a causal grip on the world, since the will can now actually affect the mechanical course of atomic events (see also Long and Sedley 1987: i. 110–12). This is not to say that all the swerves are volitionally caused, but merely that a set of alternative trajectories is available into which the atoms can be forced by volition: the will can then take advantage of this causal plasticity, and, within certain limits, mould events in its image.

Although this interpretation has some attractions, it is very doubtful whether Epicurus actually held it; and even if he did, we have no idea how it was developed. Sedley rests his case largely on a papyrus fragment in which Epicurus discusses how we should view those who dissipate their natural gifts:

242 the nature of the atoms has contributed nothing to some of their actions or the intensity of their actions and dispositions, but it is the developments (*apogegennēmena*) which themselves possess all or most of the responsibility

end p.227

for certain things of this sort. It is a result of that nature that some of their atoms move with disordered motions, but it is not the atoms at any rate . . . .[21] Thus whenever something develops which acquires some distinctness from the atoms in a manner which is genuinely distinguishable (and not simply as if from another angle), it acquires its own responsibility, which it immediately transmits to the primary natures, and makes all of it one [sc. nature],[22] as a result of which those who cannot distinguish such things in this way find themselves in severe difficulties regarding the assignment of causal responsibility. (Epicurus, *On Nature* 34.[23] 21–2 = 20b LS)

Interpretation is complicated not only by the fragmentary nature of the text, but also by Epicurus' tortured and jargon-ridden style; and the translation I give differs from Long and Sedley's in several crucial respects,[24] although I shall not detail all of them. The 'developments' (*apogegennēmena*) are presumably developments of character: the way in which initial dispositions have become moulded as a result of experience and learning. Epicurus' basic claim, then, is that it is these developed dispositions of character to which responsibility for actions should be ascribed, and not to the atoms themselves, or to the 'primary natures'.

Sedley takes the latter also to be the atoms, which is a possible reading; but it may be better to suppose that the phrase rather refers to the preliminary mental dispositions of the new-born infant prior to their environmental development.[25] The point, then, is that neither the individual atoms as such which make up the soul of the neonate (and indeed of the agent as it develops), nor even the agent's preliminary, congenital dispositions, can be held solely responsible for what the agent does.

But this position in no way requires Sedley's top-down causation. The point should rather be that the individual agent is a real thing: not really distinct from his constituent atoms in some strange, wholly emergent sense, but rather simply in the sense that, considered in and of themselves, the atoms and their properties cannot explain anything as complex as agency. Agency is, then, something which emerges upon a suitable complexity of atomic

structure, but it is not thereby *radically* emergent in Sedley's sense, such that the emergent structure defies reduction to its
end p.228

atomic constituents. It is *reducible*—it is just that it is not *eliminable* in favour of an account that only refers to atoms and their primitive properties, as the Democriteans held (§§1*a–b*: **215**).

This is the point of talking of the development 'which acquires some distinctness from the atoms in a manner which is genuinely distinguishable (and not simply as if from another angle)'; i.e. we need to think of the agent not merely as a collection of soul-atoms, but as such a collection *suitably structured*, the 'one nature' that emerges as a result of the process. But this in no way requires the radical emergence of a volitional ghost in the machine. And this story, obviously, has no direct need of the swerve: certainly not (incoherently) to constitute free action, but nor even to make room for it. But then if this is right, it is hard to see why the swerve is needed at all.

Whatever else is obscure, it is at least clear that Epicurus is here seeking to avoid what he takes to be the rationally unacceptable consequences of the sort of determinism of mind and action he ascribes to Democritus and his followers. A relatively well-preserved fragment (34. 26–30 = 20c LS) has Epicurus castigate the Democriteans for not seeing that that position is self-refuting: they 'debate this very question on the assumption that their opponent is himself responsible for talking nonsense'; but such a notion of responsibility is logically unavailable to them as a consequence of the view they seek to advance. Moreover, if the opponent alleges that his behaviour (in debating) is itself necessitated, then he will be forced into a regress.

Epicurus' point here is difficult to divine; he does not, apparently, think that all regress is itself vicious, merely that it makes the Democritean stance empty of content (see Long and Sedley 1987: ii. 108 for comments): even the Democritean takes his arguments to be persuasive; i.e. they are supposed to supply his opponent with *reasons* for modifying his views. But if one is to act upon reasons, *a fortiori* one must be able to act, i.e. be free (if there is, in Epicurus' language, some 'auxiliary element or impulse within us': 34. 29); the Democritean position is thus pragmatically self-stultifying.

If the argument is directed against determinism as such, no sophisticated compatibilist will have too much difficulty evading it (although we must wait for the Stoic Chrysippus before any self-conscious form of compatibilism is advocated in the ancient world: VII.2*a*). But it is in any case unclear if the moral Epicurus intends us to draw is, specifically, anti- deterministic: the point may rather be that, since the Democriteans officially at any rate consider macroscopic objects to be unreal (they are *merely* aggregates of atoms: §1*a*, **215**), there is no real agent there at all to be motivated by reasons. It is only if the 'developments' really do produce something distinguishable from the atoms merely considered as such that
end p.229

the ascription of reasoning, motivation, and hence of responsibility, becomes at all possible.

If this is right, the Epicurean target is less the general metaphysical thesis of determinism than it is a sort of fatalism (for the distinction between them, see VII.2*c*, 3*b*), which destroys the point of ascribing any real responsibility, causal or otherwise, to an individual. This view may be discerned in another text:

243 the man who says that everything happens of necessity (*kat' anankēn*) has no ground for criticizing the man who says that not everything happens of necessity, since according to him it too happens of necessity. (Epicurus, *Vatican Sayings* 40 = 20d LS)

To say that 'everything happens of necessity' is not just to adopt a universal determinism; rather it is to say that the individual agent makes no causal contribution to the outcomes at all. It is this latter claim which renders the practices of praise and blame, reward and punishment, ethically pointless. None the less, even if it is true that Epicurus' major beef is with fatalism rather than determinism, it is far from clear that he clearly distinguished the two (indeed, they are frequently confounded: compare the views of Alexander of Aphrodisias: X.4*c*); and he may very well have thought that the only way to avoid fatalism was by relaxing the grip of bivalence over future contingents ('otherwise everything really is outside the control of the agent': views of this sort still persist, in spite of the acuity of Chrysippus and Hume: VII.2), and hence denying determinism. But it is important to see that, even if this is right, it does not commit Epicurus to making the swerve constitute human freedom, or to allowing the emergence of a distinct, non-atomically based level of macroscopic 'downwards causation', as Sedley would have it. That line of interpretation gains limited support from Lucretius, who maintains that it is the undeniability of the real existence of independent volitions which pushes the Epicureans to adopt the position the do:

244 furthermore, if all motion is always linked, a new motion arises out of old in a fixed order, and atoms do not by their swerve make some beginning of motion to break the decrees of fate, so that cause should not follow cause from infinity, from where does free volition (*voluntas*) exist for animals throughout the world? From where, I ask, comes this volition wrested away from the fates, through which we proceed wherever each of us is led by his pleasure and likewise swerve off our motions at no fixed time or fixed region of space, but wherever the mind carries us? For without a doubt it is volition that gives these things beginnings for each of us, and it is from volition that motions are spread through the limbs. . . . Nor is it the same when we move forward impelled by a blow through another person's great strength. . . . For then it is plain that all the matter of the whole body moves and is driven against our wish, while volition reined it back through the

limbs. . . . So in the seeds too you must admit . . . that there is another cause of motion besides impacts and weight, from which this power [i.e. volition] is born in us, since we see that nothing can come to be out of nothing. . . . That the mind should not itself possess an internal necessity in all its behaviour, and be overcome, and as it were forced to suffer and be acted upon—that is brought about by a tiny swerve of the atoms at no fixed place or time. (Lucretius 2. 251–93 = 20f LS)

Volition interrupts the complete linearity of causal behaviour, and explains interruptions in motion, behavioural irregularities, and so on. It is an evident fact of animal behaviour that animals do not react to stimuli in the same way as inanimate objects. Animals clearly do, as Aristotle said, have an internal principle of motion (and rest) in the way that bricks do not.

Lucretius' account is not as lucid as it might be; he apparently thinks that because animals swerve ‘at no fixed place or time’ (that is, not in response to crude mechanical laws) the atoms must too; yet the Atomists' own concept of emergence should have protected them from that gross fallacy. Moreover, nothing Lucretius says in **244** seems directly to require a swerve, or militates against any but the most crass and fatalistic of determinisms; and that is in itself perhaps grist to the mill of the interpretation I have been pressing here. The swerve is needed only ‘to break the decrees of fate’, i.e. to avoid the logical determinism occasioned by the unrestricted applicability of the principle of bivalence; and this is the sense in which it is the source of voluntary action. Yet it must be admitted that Lucretius' remarks, both here and elsewhere, strongly suggest that he at least interpreted the Epicurean swerve as occurring within the mind- atoms, and as being directly the source of free action.

Lucretius may well have been confused, or simply have misunderstood the import (and the sophistication) of Epicurus' account.[26] His puzzling remarks about ‘seeds’ may be compared with the ‘primary natures' of **242**, and with another fragment of Epicurus:

245 from the very beginning there are seeds in us which are always directing us in various different ways towards actions and thoughts and dispositions, to a greater or lesser extent. Consequently what simply develops, this or that disposition, is at first up to us; and what flows of necessity (*kat' anankēn*) in through our pores from the environment is at one time up to us and dependent upon beliefs for which we ourselves are responsible. (*On Nature* 34. 26 = 20c LS)

That is obscure: but Epicurus, like Aristotle (IV.4*b*), apparently thought there was a time in our lives when the development of our characters really

end p.231

was up to us; although, once again, quite how this was supposed to be the case remains unclear (if it is simply because we can reflect rationally on the contents of our experience and upon our beliefs, goals, and desires, and hence contrive to affect them, such a position may be described quite independently of indeterminism).

4. Later Epicurean Scientific Epistemology

(a) The Nature of Signs

We are fortunate to possess, in the form of a damaged papyrus recovered from the ruins of Herculaneum, Philodemus of Gadara's *On Signs*.[27] While the pivotal Hellenistic debate on epistemology is beyond the remit of this study, a brief glance at *On Signs* is in order in so far as it demonstrates the extent to which the Epicureans were prepared to go in pursuit of their empiricism. We noted above that Epicurus was prepared to make use of arguments that purported to show that it was only on the basis of the atomic hypothesis that the world of phenomena could be accounted for at all—inference to the only explanation. The Stoic doctrine of the indicative sign is precisely parallel to this. Signs are supposed to reveal facts which are in some way non-evident (*adēlon*); Sextus Empiricus preserves two slightly different accounts of how, according to 'dogmatic' (i.e. non-sceptical) philosophers, things might be non-evident, and although the accounts differ on details, they agree in all essentials.[28] They are usually supposed to embody the views of the Stoics,[29] or Stoic-influenced doctors. Here is the first:

246 of matters, then, according to the Dogmatists, some are (*a*) pre-evident (*prodēlon*), some (*b*) non-evident (*adēlon*); and of the non-evident, some are (i) totally (*kathapax*) non-evident, some (ii) temporarily (*pros kairon*) non- evident, and some (iii) naturally (*phusei*) non-evident. Pre-evident are those which come to our knowledge from themselves, e.g. that it is day; totally non- evident are those which are not of a nature to fall under our knowledge, such as that the number of the stars is even; temporarily non-evident are those which, although they possess an evident nature, are now not evident to us because of certain external circumstances, as the city of Athens is to me now; while the naturally non-evident are those which do not possess a nature such

end p.232

as to fall under our evidentness, such as the intelligible pores. (Sextus Empiricus, *Outlines of Pyrrhonism* 2. 97–8)

Signs fall into two classes, since

247 the pre-evidents do not, they say, need a sign, since they are apprehended of themselves. Nor too do the totally non-evident, since they are at bottom inapprehensible. But things which are either temporarily or naturally non- evident are apprehended by means of signs, not of course the same ones, but the temporarily non-evident by way of commemorative signs and the naturally non-evident by way of indicative signs. (2. 99)

Class (*a*) are apprehended without mediation, while (*b*i) are totally unknowable; but (*b*ii) may be grasped via 'commemorative signs', and (*b*iii) by way of 'indicative' signs. The indicative sign is defined by the Stoics as

248 an antecedent proposition in a sound conditional, which is revelatory of the consequent. (2. 101)

An indicative sign brings to light a hidden fact, one which could not have been discovered simply by observation. For instance, sweating is an indicative sign, according to the Stoics, of the existence of invisible pores in the skin (2. 140; cf. Sextus, *Against the Professors* 8. 146). The Stoics construe sign-inference conditionally:

(1) [A] if there is sweating, then there are invisible pores in the skin;

(2) there is sweating;

so

(3) there are invisible pores in the skin.

The conclusion (3) is non-evident; it is inferred from (1) by way of the evident fact (2). The issue concerns the status of (1). The Stoics held it to rest upon a further argument:

(4) [B] it is not possible for liquid to pass through a solid body;
(5) sweat passes through the body;
so
(6) the body is not solid. (ibid. 8. 309)

(4) expresses, according to the Stoics, a necessary, conceptual truth, which grounds the necessity of the inference to (3). Hence the existence of pores in the skin, for the Stoics, is something that can be known with certainty. The Epicureans too allow that empirical matters admit of a certain sort of certainty: but the basis of that certainty is, for them, itself empirical. This

is the core of the dispute between the Stoics and the Empiricists as it is (imperfectly) preserved in Philodemus'*On Signs*

.

(b) Signs and Generalizations

The Stoics hold that empirical generalizations cannot form an adequate basis for scientific knowledge of an explanatory kind, since such generalizations are, of their very nature, defeasible. How can we know that future experience will continue to confirm what we have already observed? How large a sample is required? Can any sample be large enough? Moreover, inductive generalizations proceed on the basis of similarities; but what sort of similarities are relevant, and what are not? In any case, the inference from things observed to things not (yet) observed is valid only on the assumption that the property in question really is uniformly true of the class at issue: but that begs the inductive question. 'Inductive' inferences are only valid if the property in question cannot fail to hold of every member of the class (i.e. that it satisfies an argument like [B]): but then it is not an induction at all. These arguments are rehearsed by Philodemus in the first surviving part of *On Signs* (1–12).

In reply, Philodemus allows that some truths may be known in the Stoic manner (most notably that there is void: ibid. 17; §2*c*, **229** above), namely by noting the impossibility of the antecedent if the consequent of the conditional is denied (the so-called 'method of contraposition'): that is, the antecedent conflicts strongly with the denial of the consequent.[30] None the less, Philodemus argues, there is a place for empirical generalizations whose contents are certain in a different sense: the negation of the generalization is inconceivable, but not logically impossible. Indeed, in another passage he strongly suggests that the contrapositional proof of the existence of void is itself dependent upon inductive inference:

249 for we establish that if there is motion there is void not otherwise than by showing via the method of analogy [i.e. induction] that it is impossible for motion to occur without void. (ibid. 13)

That is, the necessity involved in the case of a true contrapositional proof is not logical, but empirical: it is only by experience that we come to know, for example, that solid bodies do not permit the transfer of liquid, and only then can we use that fact in the pursuit of further knowledge.

But what kinds of inductive inference are able to generate such necessity? The examples Philodemus gives concern human mortality: the Stoics

end p.234

allege that it is only by illicitly begging the question at issue that one may infer inductively that all men are mortal on the basis of the fact that all in one's experience and in recorded history are mortal. Philodemus (23, 38) replies that he does not include the conclusion in the premisses—rather

> **250** taking what is common, we infer from the fact that all men among us are alike even in being mortal, that universally all men are liable to death, since nothing opposes the inference or pulls us even a step towards the view that they do not admit of death. (ibid. 23)

The Epicureans developed an epistemology of confirmation, *epimarturēsis*, according to which a proposition received confirmation according to the extent to which different observations were seen to be in agreement with it (Epicurus, *To Herodotus* 50–1; cf. 48), and, moreover, where nothing disconfirmed it. But there are evident difficulties involved here. The fact that there has as yet been no disconfirmation of the view that all men are mortal does not show that it must be true; and hence the sense in which it is 'inconceivable' that men could be immortal seems at best to be purely epistemic. Certain passages in Philodemus suggest that the story might go as follows: we derive our conception of 'man' solely by way of our experience of individual human beings; but that concept is itself so structured as to make the notion of an immortal man inconceivable. Concept- formation is, then, a purely empirical matter; but the concepts thus formed have certain necessary entailments.[31] But then the question arises as to why, even if it is true of our concept 'man' that it precludes immortality, should we believe that our concept is the right one?

Furthermore, the Stoics pointedly ask why, if the inductive method is to be allowed, can we not argue from the destructibility of all bodies in our experience to the destructibility of all bodies (and hence to the refutation of atomism), and equally to the claim that all bodies of whatever size must be coloured: ibid. 7. Philodemus replies that it is not *qua* body that macroscopic objects are either destructible or coloured (25), and hence that no such inference is valid. But he owes us an account of how we can know which properties belong to what as such, and which are merely adventitious. This raises once again the crucial Stoic question: what counts as *relevant* similarity in the case of inductive inference? What needs to be true about predicates *F* and *G* to make inferences of the form '*a* is like *b* in respect of *F*; *a* is *G*; hence *b* is [likely to be] *G*' valid? These are large and difficult questions, the problems of induction and projectibility, which in one form or another still plague the methodology of science:[32] and it is

end p.235

clear that the Epicureans at least did not possess adequate answers to them, if indeed there are any such answers.

The Epicureans, then, leave some crucial methodological questions unresolved (we shall meet them again in the case of the Empiricist doctors: IX.2; cf. also the Eight Modes of Aenesidemus: VIII.2). But their defence of empiricism exposes a fundamental problem with the Stoics' own methodology: how can they justify the necessity of their sign-inferential conditionals otherwise than by experience? On the other hand, the Epicureans do not rest content (as Empiricist doctors and others were to do) simply with commemorative signs, cases in which the consequent of the sign- conditional is merely temporarily non-evident. They too want to penetrate to the hidden core of things, and, at least officially, believe that their theory regarding the ontological fundamentals of the universe is certain and indubitable: it is an inference not merely to the best, but, in quasi-transcendental manner, to the only explanation. Yet in view of the considerations urged by Aenesidemus (VIII.2), the Epicureans might have done better to allow the doctrine of multiple explanations a wider ambit.

5. Conclusions

Atomism was developed to answer the challenge of Parmenides (I.3*d*, 4*d*): the unalterable atoms evade Parmenides' embargo on change, while the existent nothingness of void rehabilitate motion (§1*a*). Democritus fashions a severely parsimonious ontology: all that exists are atoms and the void; everything else must be reduced, in the strongest possible way, to them (§§1*b*–*c*). Teleology is rejected: purely mechanical principles will suffice, in an infinity of time and space underwritten by the Principle of Sufficient Reason, to generate the worlds (§§1*d*–*e*).
The ontology of Democritus' atomist descendant Epicurus was more accommodating: he allowed that perceptual properties were real, although emergent upon their atomic foundations (§2*b*); and he allowed that, while in general there were no uncaused events (§2*c*), each atom was none the less prone to swerve minimally from its trajectory for no reason (§§2*a*–*d*), in order to allow for the generation of worlds and (perhaps) for human freedom (§§3*a*–*b*). Epicurus follows Democritus in rejecting teleological explanation in nature altogether (§2*e*); and develops the Theophrastean view (V.3*b*) that in some cases it is appropriate to offer a variety of equally plausible, competing ‘explanations’ of physical phenomena, where the evidence is unable to discriminate between them (§2*f*). In epistemology,
end p.236

the Epicureans reject the Stoic view that fundamental physical principles with empirical content can (at any rate in general) have the status of logically necessary truths, although they can be inductively secure (and inductively secured) *conceptual* necessities (§§4*a*–*b*).
end p.237

VII The Stoics

R. J. Hankinson

1. Stoic Causal Theory

(a) Background

When Aristotle used the terms *aition* and *aitia*, he was concerned with explanation in a broad sense. By contrast, the philosophers of the Hellenistic period tend to restrict their scope to action, production, and agency. We may briefly take stock of what the various Hellenistic theorists of cause and explanation tended to have in common with each other, and how their shared presuppositions differed from those of their predecessors. Hellenistic causal theorists were materialists. Their materialisms differed considerably from one another, but there is little or no dispute among them that causation is corporeal, and transmitted by contact. Sextus Empiricus remarks that ‘some people say that body is what can act and be acted upon’ (*Outlines of Pyrrhonism* 3. 38; cf. *Against the Professors* 9. 366), defining corporeality in terms of causal efficacy; Zeno of Citium (344–262 bc), the founder of Stoicism, insists (against the Platonist Xenocrates: X.1*b*) that incorporeal substance could have no power (Cicero, *Academica* 1. 39). Moreover, the Stoics' insistence upon the active nature of causation (cf. § *f*, **265** below) has the consequence that design and purpose disappear from the causal vocabulary.[1] This is not to say that they vanish completely, or that they play no role in explanation; but for something to be properly a cause, an *aition*, it must now be more than just an essential component in a complete explanation: it must be dynamic, something which actually does something. When Sextus Empiricus remarks that ‘some say that cause is corporeal, others that it is incorporeal’ (*Outlines of Pyrrhonism* 3. 14), the proponents of causal incorporeality he presumably has in mind are Platonists and Pythagoreans (cf. *Against the Professors* 9. 364); at any event, he feels entitled to say that

251 in general it would appear that in their view [i.e. the view of the Dogmatists as a whole] a cause is that because of whose activity an effect comes about. (Sextus Empiricus, *Outlines of Pyrrhonism* 3. 14)

end p.238

Moreover, it is generally agreed that causation is universal. Galen goes so far as to treat propositions like ‘nothing occurs without a cause’ and ‘nothing comes to be from nothing’ as a priori metaphysical axioms (*On the Therapeutic Method* 10. 36–7: XI.2*b*; cf. Alexander of Aphrodisias, *On Fate* 22. 192. 14; *Mantissa* 22: X.4*c*, **385**). But even if every event has a cause (a thesis explicitly rejected by the Epicureans with their swerve: VI.2*a*) it does not yet follow that the world is deterministic in character: for that to be the case every event must have a uniquely determining cause. But that is a thesis that the Stoics were prepared to hold:

252 there are, then, many types of cause,[2] but they say that it is equally true of all of them that it is impossible, when all the circumstances surrounding both the cause and that of which it is the case are the same, for things sometimes to turn out in a certain way and at others times not; for there would be, if this were to happen, an uncaused motion. (Alexander, *On Fate* 22. 192. 21–5; cf. VI.2*a*, **226**)

(b) Stoic Materialism

Stoic physical doctrine may be briefly summarized. Their physics distinguishes fundamentally between the principles of the Active and of the Passive. This too, Sextus thinks, is relatively uncontroversial: 'it is agreed by most that of principles (*archai*) some are material and some efficient' (*Outlines of Pyrrhonism* 3. 1); and his language is designed to underline the agreement here between Aristotle and the Stoics. However, the Active and the Passive are, for the Stoics, material in nature, the light, volatile elements of air and fire (associated respectively with the qualities cold and hot) supplying the active power, while the grosser (moist) water and (cold) earth form the passive substrate (the principles are, however, eternal, unlike the elements which are susceptible of the familiar transmutations: Diogenes, *Lives* 7. 134–7). In particular it is the dynamic compound of air and fire known as *pneuma* which permeates all things, and is responsible not merely for the intelligence of the universe but also for the cohesion of its component objects, literally holding them together by dint of its internal tension.
This permeation is, for the Stoics, absolute, through and through. *Pneuma* is not to be conceived of as a kind of internal skeleton for things, spatially distinguished from that which it permeates: rather the inter- penetration in the Stoic continuum is an Anaxagorean total intermixture (**254** below; cf. I.4*b*). This picture was much derided by other theorists: the Peripatetic Alexander of Aphrodisias, for example, attacks it on the grounds that it subverts common conceptions, as well as relying on a
end p.239

confusion between formal and material causes (*On Mixture* 9–11: see further X.4*b*). For the Stoics, *pneuma* is in a constant state of internal, dynamic tension, which manifests itself in a constant backwards and forwards movement (ibid. 11. 224. 23–6 Bruns;[3] Nemesius, *Nature of Man* 70–1), a condition which is somehow supposed to be responsible for its overall stability. Galen, unlike Alexander and the Peripatetics, does not reject this suggestion out of hand: he notes that in the case of a limb which is extended but motionless, there is an equal balancing of continuously acting forces upon it (*On Muscular Movement* 4. 401–2); and he continues:

253 imagine a bird on high which seems to rest in the same place. Should it be said to be motionless, as though it were suspended from above, or to be moving with an upward movement to the same extent as its body's weight drags it downwards? The latter seems to me the truer description; for if deprived of its life, or even its muscular tension, you would see it falling rapidly to earth, from which it is clear that the natural downward tendency of its body's weight was equally balanced by its upward movement in respect of its soul's tension. (Galen, *On Muscular Movement* 4. 402–3[4] = *SVF* 2. 450 = 47k LS)

Furthermore, there is evidence that the Stoics conceived of this internal dynamic tension, which literally bound all things together (and was the foundation for their belief in *sumpatheia*, the universal interconnection of all things: cf. §2*d*; VIII.5*a*), as operating in the form of waves: at all events, they used the behaviour of waves as an analogy for the propagation of causal influence in a continuum (Aëtius 4. 19. 4).

Whatever the actual mechanics of this are supposed to be,[5] the Stoics clearly postulate active causes not only for events and processes, but also for persistent material states of affairs, which they labelled *aitia sunektika*, or 'containing causes':[6]

254 if after they had been mixed these ingredients (*pneuma*, water, and earth) were all to remain in their original state, that would imply that their minute parts had simply been juxtaposed and not that they had been totally intermingled. But this is just what Empedocles thought. For he used to hold

end p.240

the view that natural bodies are produced not by an intermixture of the four elements but by their combination and for this reason on this point his theory coincides with that of Epicurus and Democritus, whereas neither the Stoic philosophers nor Aristotle talked of juxtaposition. (Galen, *On Containing Causes* 5. 2–3).

Galen (like Alexander) rejects as incoherent the view that every body requires an actual active containing cause:

255 if every single existent thing requires a containing cause without which it cannot exist, that cause, as it is an existent, must inevitably have another containing cause itself, which must in turn have yet another—and so on *ad infinitum*. (ibid. 6. 3; cf. Galen, *On Repletion* 7. 524–8; Alexander, *On Mixture* 11. 224. 13)

In other words, if

(1) every cause is corporeal,

and

(2) every corporeal thing requires some further distinct containing cause of its existence,

then if there is anything at all, there must exist infinitely many bodies organized in an endless hierarchy. That supposition is not perhaps incoherent: but it is at the very least ontologically extravagant. It is not clear, however, that the Stoics need be committed to it; they can deny (2) in its full generality, if they allow *pneuma* to be self-cohesive in spite of its corporeal nature. But then they owe us an explanation of how the lighter and more easily dispersed elements can hold together the heavier and more stable ones, a position which Galen at least takes to be absurd. Solid objects are cohesive, Galen thinks, just because that is what solidity amounts to: 'the fact that they (i.e. rocks and metals) are solid depends on this very quality, namely their self-coherence' (*On Containing Causes* 6. 5).

But even so, there is no logical bar to treating solidity as a derived property, to be explained reductively in terms of other more basic properties, in the manner of contemporary physical chemistry. Moreover, the requirement that objects require a cause for their persistence does not conflict with the requirement that a cause be active: constant, active pneumatic tension is needed to prevent solid objects like cups and tables from collapsing into formless heaps of matter; and these containing causes, although intimately intermixed with the passive constitutive material, are still distinguishable from them in analysis. In this sense Frede (1980: 243) rightly calls containing causes ‘the Stoic analogue to Aristotle's formal cause’.
end p.241

(b) The Stoic Analysis of Causation

256 (i) Zeno says that a cause (*aition*) is ‘that because of which’, while that of which it is a cause is an attribute; and that the cause is a body, while that of which it is a cause is a predicate. (ii) He says that it is impossible that the cause be present yet that of which it is the cause not belong. (iii) This thesis has the following force. A cause is that because of which something occurs, as, for example, it is because of prudence that being prudent occurs, because of soul that being alive occurs. . . . (iv) Chrysippus says that a cause is ‘that because of which’; and that the cause is an existent or a body. . . . (v) He says that an explanation (*aitia*) is the statement of a cause, or the statement of a cause *qua* cause. (Stobaeus, *Anthology* 1. 13. 1c = *SVF* 1. 89, 2. 336 = 55a LS)

(i) suggests that the proper form of a causal statement is

(C1) x causes F,

where x picks out an object and F an attribute. C1 is clearly inadequate, and may be supplemented by way of a remark in Sextus:

257 the Stoics say that every cause is a body which becomes a cause to a body of something incorporeal. For instance the scalpel, a body, becomes the cause to the flesh, a body, of the incorporeal predicate ‘being cut’. (*Against the Professors* 9. 211 = *SVF* 2. 341 = 55b LS)

C1 may be thus amended to

(C1*a*) x causes y to be F;

causation is a triadic relation that holds between two bodies and an incorporeal attribute,[7] and occurs when one object effects a change in the condition of another.

256(ii) expresses a thesis of causal sufficiency:

(C2)$(x)(y)$(if x causes y, then whenever x occurs, y occurs),

and implies further that cause and effect must be coeval, i.e.:

(C3)$(x)(y)$(if x causes y, then whenever x is present, y is present).

Thus, apparently a cause must invariably produce, and always be producing, its effect. In that apparently innocent contention lurks the seed of the Sceptics' assault on the coherence of the Dogmatic concept of causation (VIII.3).

The condition of the contemporaneity of cause and effect is clearly satisfied by the original concept of an *aition sunektikon*, a cause continuously operating to preserve the material integrity of an object (§ *b*, **254**).
end p.242

But clearly not everything which satisfies C2 and C3 need be a conserving cause:

258(i) the majority of them hold that of the causes some are containing (*sunektika*), some co-operative (*sunaitia*), and some auxiliary (*sunerga*); (ii) and that causes are containing if, when they are present the effect is present, when they are removed the effect is removed, and when they are decreased the effect is decreased (thus they say that the application of the noose is the cause of the strangling); (iii) and that a co-operative cause is one which contributes a force equal to that of its fellow cause to the occurrence of the effect (thus they say that each of the oxen drawing the plough is a co-operative cause of the drawing of the plough); (iv) and that an auxiliary cause is one which contributes a slight force to the easy production of the effect, as for instance when two men are lifting a heavy weight with difficulty, a third appears to lighten it. (Sextus Empiricus, *Outlines of Pyrrhonism* 3. 15)[8]

The notion of a *sunaition* to be found in Sextus and elsewhere[9] does not correspond with Plato's usage (III.4*b*; cf. X.1*b***354** and n. 10). Two (or more: Sextus' account is readily generalizable to cover any number of cooperative causes) *sunaitia* are the co-operative causes for an effect when each of them supplies some of the causal power required for it, by cooperating to bring about an effect which neither could achieve singly; they are thus collectively equivalent to an *aition sunektikon*.[10] A *sunergon*, by contrast, is an item which contributes to an outcome by reinforcing the effect of an *aition sunektikon* without itself being capable of bringing about the effect on its own.[11] So Clement:

259whereas the auxiliary cause aids the containing cause, so as to intensify what comes about through the latter, the co-operative cause does not correspond to the same conception, since a co-operative cause may exist where there is no containing cause. For the co-operative cause is conceived along with another which is itself equally incapable of producing the effect on its own, since they
end p.243

are causes co-operatively. The difference between the co-operative cause and the auxiliary cause lies in the fact that the co-operative cause produces the effect along with another cause which is not independently producing it, whereas the auxiliary cause, in creating the effect not independently but by adding to another, is acting as auxiliary to the very cause which is independently creating the effect, so that the effect is intensified. (Clement of Alexandria, *Miscellanies* 8. 9. 33 = *SVF* 2. 351 = 55i LS)

258(ii) establishes that containing causation involves functional dependence:[12] *aitia sunektika* are co-variant as well as coeval with their effects (cf. ps.-Galen, *Medical Definitions* 19. 393), which requires a further strengthening of C3. Containing causes will now conform to the following definition:

(C4)(x)(y)(if x causes y, then whenever x is present, y is present, and any variation in the strength of x is matched by an equivalent variation in the strength of y).

These containing causes, unlike those Stobaeus ascribes to Zeno in **256**(iii), are not merely conservative: they are genuine agents of change.

As it developed, then, the concept of containing causation ranged beyond its initial role in the explanation of persistent states.[13] In so doing it became clear that if C2's thesis of causal sufficiency was to be retained, the concept was in need of further refinement, since frequently no one individual item is exclusively responsible for a given effect: consider the team of oxen ploughing. Thus the requirement that there be a unique cause for each event is relaxed, although there must still be some collection of causal factors, the *sunaitia*, or co-operative causes, which jointly meet the conditions of both sufficiency and contemporaneity (**258**(i), **259**).

(d) Antecedent Causes

Even so, many items are ordinarily taken to be causes which fail to satisfy C2–4. Causes are often thought to precede their effects; and they are not invariably supposed to be sufficient conditions:

260 some of them, however, have said that things present can be causes of things future as well, as antecedents (*prokatarktika*), as for instance protracted exposure to the sun of fever. But some reject this since the cause is relative, and

end p.244

relative to the effect, and hence cannot precede it as cause.[14] (Sextus Empiricus, *Outlines of Pyrrhonism* 3. 16)

The Stoics, then, allow that there can be other kinds of cause besides *aitia sunektika*, and hence that C2–4 do not exhaustively capture the concept of causation. Chrysippus (*c*.280–206 bc), the Stoics' third leader, turned to antecedent causes to ground his account of human freedom (§2*c* below; as **260** suggests, it also found a ready home in medical contexts: XI.2*b*–*c*). An antecedent cause precedes its effect, and consequently also the containing cause of that effect:

261 when antecedent causes are removed the effect remains, whereas a containing cause is one during whose presence the effect remains and on whose removal the effect is removed. The containing cause is called synonymously the perfect (*autoteles*) cause since it is self-sufficiently productive of the effect. (Clement, *Miscellanies* 8. 9. 33 = *SVF* 2. 351 = 55i LS)

Chrysippus (Cicero, *On Fate* 42–3; Aulus Gellius, *Attic Nights* 7. 2. 10) likened human action to a cylinder's rolling. It needs some external stimulus, analogous to the perceptual impressions which initiate human activity, to get it moving; but equally for it to be able to roll at all it must have a certain structure (i.e. its surfaces must be curved): its rollability is a function of its internal constitution. The initial stimulus is, in contrast with the persisting disposition to roll, an ephemeral, particular, and external event; and ancient accounts of antecedent causes are unanimous that they are external to the things of which they are causes (cf. *On Fate* 24).[15] Here is Cicero's report of Chrysippus' account of causation:

262 ‘some causes’, he says, ‘are perfect (*perfectae*) and principal (*principales*), others are auxiliary (*adiuvantes*) and proximate (*proximae*). Hence when we say that everything takes place by fate from antecedent causes, we should not be taken to mean by perfect and principal causes, but by auxiliary and proximate causes’. Accordingly he counters the argument which I have just set out as follows: ‘if all things come about by fate it does follow that all things come about from prior (*antepositae*) causes, but not from principal and perfect but from auxiliary and proximate causes’. (Cicero, *On Fate* 41 = *SVF* 2. 974 = 62c LS)

Cicero's account is puzzling in many ways. He lumps ‘auxiliary and proximate’ causes together, while explicitly distinguishing them from their

end p.245

‘perfect and principal’ counterparts.[16] However, at *On Fate* 44, albeit in his own voice, Cicero refers to ‘proxima illa et continens causa’, where ‘continens’ must translate *sunektikon*.[17] *Aition proseches* is the probable original for ‘causa proxima’;[18] and Galen (*On the Causes of Pulses* 9. 107) insists that *proseches* and *sunektikon* are synonyms. But that would make nonsense of **262**, whose proximate causes must be antecedent causes. Moreover, Galen reports (*On Antecedent Causes* 14. 174–6) that for the third-century physician Erasistratus an *aition proseches* was a genuine cause of something, rather than merely a causally relevant antecedent,[19] and was proximate in the sense of being immediately contiguous to its effects (see further IX.2*a*; XI.2*b*). *Aitia prosechē* are thus in this sense proximate causes, and may be identifiable with *aitia sunektika*; but there is no evidence that Chrysippus ever adopted this usage, and Cicero is probably wrong to identify proximate and containing causes in his exegesis of Stoic doctrine.

More important is the question of whether perfect and principal causes are supposed to be different items, or whether they are rather simply alternative designations of the same thing (and the same goes for proximate and auxiliary causes).[20] ‘Principalis’ may be Cicero's translation of *kurioōtaton* (‘most important’), an epithet sometimes (although apparently non-technically) applied to causes; if so, it would be indistinguishable in meaning from ‘perfecta’. Yet it might also be, as a result of another linguistic confusion, that Cicero's ‘principalis’ renders *proeēgoumenon*. Many texts talk of *aitia proēgoumena*, preceding causes; but most of them are medical, and none appear to be authentically Stoic.

Alexander of Aphrodisias, who uses the term in his rebuttal of Stoic determinism, apparently wavers between the senses ‘primary’ and ‘preceding’.[21] Alexander attacks the Stoic conception of chance as what is epistemically uncertain by pointing out that ordinary people describe events as being chance ‘when they supervene on the *aitia proēgoumena* of other things’ (*On Fate* 8. 172. 17–23): here ‘*proeēgoumenon*’ ought to mean ‘principal’ or ‘primary’—in fact, a synonym for what Alexander elsewhere

end p.246

(following Aristotle) calls ‘*per se* causes’ (*Mantissa* 24; see X.4*b*; cf. Galen: XI.4*b*, **402**). Later, however, Alexander challenges the Stoics to explain how anything can be a matter of chance if this common intuition is correct, and yet ‘everything that is or comes to be does so of necessity from certain prior and preceding causes’ (*On Fate* 8. 173. 13–21):

here precedence rather than primacy seems to be important (and Alexander may be deliberately distorting the Stoic position here for polemical purposes).[22] Alexander's evidence is thus equivocal: if *proēgoumenon* means 'primary' (and if *proēgoumenon* is indeed the word Cicero renders as 'principalis'), once again it will simply reduplicate 'perfecta'. But if it bears the sense of 'preceding', then, since perfect causes are equivalent to *aitia sunektika* which are coeval with their effects, Cicero's perfect and principal causes would have to differ in denotation. On the other hand, Alexander's 'preceding causes' here at least ought to be antecedent causes, since it is they which are, for the Stoics, the instruments of fate (see §§2*b–c* below). The case is murky indeed.

(e) Dispositions and Powers

If 'perfecta' and 'principalis' really do pick out different parts of the causal structures, one might cash out the distinction in the following way. Even when stationary, Chrysippus' cylinder still possesses, in virtue of its form, its intrinsic dispositional rollability. Upon receipt of a suitable external stimulus (and other things being equal) that potentiality will be actualized; and it is obviously helpful from the point of view of rigour to distinguish between pure dispositional properties and their actualized states (not least in the case of psychology, Chrysippus' concern in **262**).

Thus we might suppose that Chrysippus did indeed use *aition proēg- oumenon* to refer to the persistent dispositional conditions of an agent in virtue of which a particular external stimulus would produce a particular result. Clement offers an example:

263 of causes, some are antecedent (*prokatarktika*), some containing (*sunektika*), some auxiliary (*sunerga*), some prerequisite (*hōn ouk aneu*). Antecedent are those causes which primarily provide the impulse towards the coming to be of something, as beauty is to those intemperate in love; for when it is seen by them it conditions the erotic disposition, but not however in such a way as to necessitate it. (Clement, *Miscellanies* 8. 9. 25 = *SVF* 2. 345)

The last clause is vital: even when the disposition has been primed by a suitable stimulus, action in accordance with that disposition is still not inevitable. The agent still needs to assent to the impulse which the

end p.247

disposition, roused to action by the impression, provides.[23] But Chrysippus' rolling cylinder analogy provides for only two types of internal state, namely the underlying potential for action (its morphological suitability to roll), and its stimulus-induced actualization (its rolling): nothing in the illustration corresponds to the additional component (of human actions at least) of assent to the content of the aroused disposition. If assent is already a part of Chrysippus' psychophysical causal machinery, his illustration seems not to be a particularly happy one. But in the final analysis, it seems plausible to suppose that Chrysippus did indeed call the aroused disposition the *aition proēgoumenon*, which upon receipt of assent becomes the perfect cause of action.[24]

Let us turn now to consider the status of the other pair of appellations, ‘adiuvans’ and ‘proxima’. ‘Adiuvans’ almost certainly renders *sunergon*; and, on the supposition that ‘proxima’ at least denotes *prokatarktikon*, it remains to determine their relationship. Frede (1980: 240–1) treats them as effective synonyms, then wonders ‘how an antecedent cause can be conceived of as a *sunergon*, if a *sunergon* is the kind of item which helps to bring about the effect by making it easier’. Frede argues that there emerged in the Hellenistic period two distinct causal taxonomic triads: (A) *autoteles*, *sunaition*, and *sunergon* on the one hand, and (B) *sunektikon*, *proeēgoumenon*, and *prokatarktikon* on the other. Although their components are not uniformly separated (*sunaitia* and *sunektika* often occur together: **258–9**), Galen significantly never uses the terms of (A) at all. Yet even if Frede is broadly right, (A) and (B) can hardly be equivalent. (A) distinguishes different types of contemporaneous cause, while (B) explicitly introduces temporal differences: thus they must fulfil distinct (if complementary) roles.

Thus Frede's attempt to make *aitia prokatarktika* part of the class of *sunerga* seems mistaken, and there is no reason to think that Cicero's ‘adiuvans’ and ‘proxima’ should be taken to be synonymous. The *causa proxima* for the cylinder's movement is the initial push which gets it going as distinct from the internal property in virtue of which it continues to do so, a property significantly described by Cicero as a *vis*, or force. The cylinder, once set in motion, goes on rolling ‘suapte vi et natura’, by its own force and nature. The cylinder's decisive properties, then, involve more than merely a passive suitability for being rolled: they include an internal force which is actually responsible for the moving, for keeping the process going.[25]

end p.248

Even so, something may still intervene from the outside to assist the process, by giving the cylinder an extra push: this, then, would be an auxiliary cause, a *sunergon* for the process, being neither essential for keeping the rolling going, nor part of the internal cause which does so. Antecedent and auxiliary causes are thus distinct items, although they have key features in common (they are external stimuli; and they are not sufficient for their effects).

There are, then, philosophical reasons for Chrysippus to assign distinct roles to each member of his two pairs. He sought to distinguish between an object's internal propensity for action, and that propensity in a galvanized state (respectively the *aition proeēgoumenon* and *autoteles*); and he differentiated between the initial triggering of the disposition to establish a perfect cause and subsequent assistance to it. Both of these distinctions were taken over and modified by later theorists, philosophical and medical (cf. XI.2).

(f) Causes and Conditions

The notion of an internal cause is still in need of clarification:

264 they [sc. the Stoics] say that there is a difference between whether a thing is of such a kind that something cannot be brought about without it, or such that something must necessarily be brought about by it. None of the causes mentioned therefore [namely a set of remote prerequisite conditions] is really a cause, since none by its own force brings about that of which it is said to be the cause; nor is that which is a condition of a thing's being brought about a cause, but that which is such that when it is present that of which it is the cause necessarily is brought about. (Cicero, *On Fate* 36 = *SVF* 2. 987; cf. **363**)

The distinction between causes proper and mere prerequisites was already made by Plato (III.1a), and we will meet it again in Galen (XI.3*c*; cf. X.2*a*). Clement also adverts to it in **263**, and gives an example:

265 all of the causes can be exhibited in order in the case of the pupil. The father is the antecedent cause of learning, the teacher the containing cause, the pupil's nature the auxiliary cause, while the time holds the place of the prerequisite. (Clement, *Miscellanies* 8. 9. 25 = *SVF* 2. 345)

With the best will in the world on both sides, the teacher cannot teach, nor can the learner learn, unless there is time available to do it in. But the time does not actively bring the learning about (compare Galen's attitude to

end p.249

place: XI.3*c*, **404**). Even so, Clement concedes that 'many say that what does not hinder is a cause as well': however,

266 we say against them that the conception of cause involves doing, being active, and performing. In this respect at least that which does not hinder is inactive. (ibid. 1. 17. 83 = *SVF* 2. 353)

Thus people like the eclectic Potamo of Alexandria who make place into a causal category are mistaken (Diogenes, *Lives* 1. 21; cf. XI.4*c*).[26] This Stoic insistence on the primacy of agency allows Seneca, in spite of the terminological proliferation we have examined, to contrast the unity of the Stoic position with the 'crowd of causes' offered by the Peripatetics and others (*Epistles* 65. 11 = *SVF* 365a: X.2*d*).

Properly analysed, then, causal sequences are streams of activity: but to explain that activity one must point to the internal conditions of the causal objects themselves. Moreover, if fate is indeed an uninterrupted sequence of causes,[27] then (given some plausible assumptions) it is hard to evade determinism: and the Stoics did not try to. But if the unfolding of events is determined from all eternity, how can we hold agents responsible for their actions? The Stoics reply that the ascription of responsibility is justified since agents do what they do because of their own dispositions (**264**; cf. Epicurus: VI.3*b*, **242**). Cicero reinforces the point:

267 'cause' is not to be understood in such a way as to make what precedes a thing the cause of that thing, but what precedes it effectively: the cause of my playing ball was not my going down to the campus, nor did Hecuba's giving birth to Alexander make her the cause of the death of Trojans, nor was Tyndareus the cause of Agamemnon's death because he was the father of Clytemnestra. (Cicero, *On Fate* 34).

Cicero quotes Ennius' version of the opening of Euripides' *Medea*: would that trees had never been felled on Pelion to build the Argo (since if they had not, Jason would never

have sailed to Colchis). But Cicero points out that, while these causal antecedents are perhaps necessary conditions of Medea's falling in love with Jason, her rejection, jealousy, and hence her eventual filicide, they cannot be morally responsible for it, since they are not properly its causes. Clement makes a similar point using a similar example at *Miscellanies* 8. 9. 27: all of the preceding events are chance
end p.250

causes (and hence not real causes) of Medea's crime: they are not such as to determine it (cf. **363**). Rather facts about the agents' characters are the genuine causes of their being disposed to act as they do, although their actually acting thus is itself precipitated by perceptual information coming from the outside. The fundamental Chrysippean distinction between internal states and external stimuli is invoked once again.

(g) Causes and Time

Thus the Stoics seek to accommodate two basic intuitions about causes. In one sense (captured by C2), causes are contemporary with their effects; but equally, if the causal structure of the world unfolds diachronically, in some (distinct) sense causes must precede their effects. Sextus, characteristically, takes this to show the incoherence of the Dogmatists' notion of causing (*Outlines of Pyrrhonism* 3. 25–30: VIII.3*b*, **300–1**). But while that conclusion is too quick, the temporal issues involved are both important and tricky. If causes are sufficient (in the circumstances) for their effects, as any determinist must hold, why does causing take time? Why does the whole causal history of the world not collapse into an instant? There is no actual evidence of the Stoics grappling directly with this profound problem; but their idea that the fundamental causal facts about the world are the internal properties of objects is well adapted to cope with it (as indeed is Aristotle's concept of the gradual realization of form).

For the Stoics, processes (and enduring states) are ontologically basic: events are simply, in a modern jargon, three-dimensional slices through a four-dimensional continuum. And processes have temporal extension built into them. In fact, ancient texts very rarely deal with instantaneous events at all (Epicurean atomic collisions apart). Processes tend towards states which are their natural outcome (so of course Aristotle). In such an ontology, no such temporal collapse of events threatens.

This ontology emerges from disputes about the semantics of causal language:

268 the sun, or the sun's heat, is the cause of the wax being melted or the melting of the wax. For even on this they are in dispute, some saying that causes are causes of nouns like 'the melting', others of predicates like 'being melted'. (Sextus Empiricus, *Outlines of Pyrrhonism* 3. 14)

Clement agrees:

269 some are causes of predicates, e.g. 'is cut' whose case is 'being cut', others of propositions, e.g. 'a ship is built', whose case this time is 'a ship's being built'. (*Miscellanies* 8. 9. 26 = 55c LS)

The predicate-expressions are clearly suited to the denotation of states ('the wax's being melted'), while the noun-phrases naturally designate
end p.251

processes. And nothing stands in the way of adopting both modes of expression.

This suggests the following. The world contains individual things, whose status as members of natural kinds involves their possession of certain determinate attributes. These attributes are dispositional properties in virtue of which they behave in determinable ways. Those dispositions, however, are only roused into actuality by the individuals' coming into perceptual contact with certain external facts. Those episodes of contact are in a sense (a sense perfectly capturable by Aristotle's analysis) matters of chance, since the internal structures of the kinds in question do not determine them. However, for all that, the unfolding of the world, as a sequence of actualizations of causal processes in substances consequent upon their interaction, may be perfectly determinate. It is precisely this distinction upon which the Stoics rely in their attempt to make room for human freedom and responsibility.

2. Freedom and Determinism

(a) Stoic Compatibilism

Aristotle made the notion of choice central to his account of human responsibility (V.4*b*; cf. X.4*c*). We act freely, and hence are responsible for what we do, just in so far as our actions are the result of choice. If this choice is chimerical, so too will be human freedom; and determinism threatens just this result. If the universe's history is ineluctably settled for all eternity, and the unfolding of that history is the inevitable result of fate, genuine human agency, in the sense of our being able to make a difference to things, seems to be ruled out. We are simply marionettes on the strings of fate. Both determinist and indeterminist have relied upon that inference from the ineluctability of fate to the impossibility of freedom: some (such as the Epicureans: VI.3a) have argued contrapositively from the unthinkability of human bondage, while others have found the plausibility of the thesis of universal causation enough to make them deny the possibility of human autonomy, Gorgias being only the first of many. But not everyone has accepted that inference: some (pre-eminently David Hume) have tried to reconcile determinism with some account of human freedom; and Chrysippus was perhaps the first to do so.
In the last chapter, we briefly reviewed the impact of arguments owed ultimately to Aristotle that sought to relate logical and metaphysical conclusions. Aristotle's near-contemporary Diodorus Cronus elaborated his ‘master argument’ to show that, if (i) the past was necessary, and no (ii) impossibility could follow from a possibility, then (iii) the future was
end p.252

necessary and ineluctably settled as well.[28] The Epicureans countered these moves by denying that the principle of bivalence applied to future contingent propositions (VI.3*a*). The Stoics too felt obliged to take issue with Diodorus; but while they agreed in rejecting its conclusion, they disagreed as to how to do so. Cleanthes, Zeno's successor as head of the Stoa, rejected (i); Chrysippus rather more remarkably repudiated (ii), maintaining that a possibility could entail an impossibility (Epictetus, *Discourses* 2. 19. 1–5).[29]

Yet, for all their desire to evade the Diodoran conclusion that all truths are necessary, the Stoics definitely accept universal causality. Numerous texts attest to their concept of an all-embracing ineluctable fate, which they identify with the will of Zeus, Zeus himself, the *logos* of the world, among other things (cf. §§ *b*, 3*a–b* below). Indeed

270 Chrysippus the Stoic said that there was no difference between what was necessitated and fate, saying that fate was an eternal, continuous, and ordered movement. (Theodoretus 6. 14 = *SVF* 2. 916)[30]

Fate is described as a chain (or rather a rope)[31] of causes (*SVF* 2. 915, 917, 920, etc.), unravelling in an inevitably determined manner (cf. *SVF* 2. 202, 528, 913–14, 917–18, 923–4, etc.). And Chrysippus is said to have written the first book of his *On Fate* with the express purpose of demonstrating 'that everything is encompassed by necessity and fate' (Diogenianus, in Eusebius, *Preparation for the Gospel* 6. 261c = *SVF* 2. 925). Clearly, in view of this metaphysics, rejection of the Diodoran conclusion will be a tricky business. Chrysippus significantly says (in **270**) that things are *necessitated* by fate, not that they are *necessary*: and this distinction is not trivial. The Stoic account of necessity and possibility is of critical importance; and the interpretation of these modal concepts was already a subject of hot philosophical debate among Diodorus' circle. For Diodorus, necessity was simply a matter of eternal truth;[32] but the Stoics held a broader view:

271 a proposition is possible which admits of being true, there being no external factor to prevent its being true. . . . Necessary is that which, being true, does not admit of being false, or if it does so admit is prevented from being false by external factors. (Diogenes Laertius, *Lives of the Philosophers* 7. 75)

end p.253

The examples Diogenes gives are not very helpful, and the interpretation of **271** is controversial.[33] But the Stoics apparently treat as necessary things which as a matter of fact have turned out to be true, and whose truth is now unassailable (recall that Chrysippus at least accepted the first premiss of the Master argument). Even so, there will be a complete causal explanation for any such fact: they will not be matters of mere chance (indeed the Stoics interpreted the concept of chance purely epistemically, as 'a cause unclear to us': Alexander, *On Fate* 7. 172. 12, 8. 173. 13, 174. 1; *SVF* 2. 965–71; see further 2*d* below; X.4*b–c*). Moreover, the Stoics may perhaps still distinguish between necessary and contingent truths in the following manner: a proposition we would treat as contingent will only be necessary for the Stoics if there is some actual causal factor operating now to prevent its failing to be true.

The Stoics can thus evade the following objection: given the Stoic defin- ition of necessity anything that will turn out true must be necessary, not simply as a matter of logic, but because, given the adamantine necessity of fate, the world now contains the causes why things will turn out in this way: determinism is indifferent to time. The Stoics will here invoke their concept of the perfect cause (§1*d* above). For *A* to be a perfect cause of *B*, *A* must actually be acting to bring *B* about; and only when the perfect cause of some state or process is in operation is that state or process necessary.

Two distinct types of modality emerge from this. One might be labelled ‘species possibility’: some predicate *P* is species-possibly applicable to an individual of natural kind *K* just in case *K*s can, other things being equal, be *P*s. Thus Philo of Megara held that a piece of wood sitting on the seabed was still capable of being burnt, just because wood is by nature combustible (Alexander, *On ‘Prior Analytics’* 184). But there is in addition something we might call ‘actual possibility’: the submerged wood is not *actually* combustible because of the particular circumstances in which it finds itself. The Stoics will then restrict non-actual species-possibilities to future cases; but they will still allow some of them. Philo's account of possibility will apply, for the Stoics, only to the future; for the present or the past, the actual prevention condition of **271** will rule it out.

If this interpretation is right, it is false to say, as many commentators have done, that Stoic possibility must be purely epistemic in form. In Aristotle's example (*On Interpretation* 9. 19^{a}12–14), a new coat may simply wear out with use, or it might be deliberately destroyed. *Sub specie aeternitatis* there is only one thing that can happen to it on the Stoic view: fate's unravelling will see to that. However, there is nothing *now* in the

end p.254

world that prevents either outcome, for no causally efficient state of affairs is *now* making it the case that it will (or will not) be deliberately destroyed. Chrysippus' insistence that fate is an ineluctable chain of *antecedent* causes now begins to make sense.

The Chrysippean Notion of Fate

272 Between the two views held by the old philosophers, one being that of those who held that everything takes place by fate in the sense that fate exercises the force of necessity—the view of Democritus, Heraclitus, Empedocles, and Aristotle—the other that of those who said that the movements of the mind are voluntary and not at all controlled by fate, Chrysippus stands as an honorary arbiter and wished to strike a mean between the two; though he leans rather towards those who hold that mind is free from all necessity of motion . . . none the less he slips into such difficulties that against his will he lends support to the necessity of fate. (Cicero, *On Fate* 39)

Chrysippus, on Cicero's account, finds himself caught on the horns of a dilemma. Unwilling to abandon the intuition that human beings are responsible for their own actions (cf. VI.3*b*), he still wants to maintain universal causal determinism. In order to evade the dilemma, he seeks to prise assent, *sunkatathesis*, the mental act of affirming and inclining towards the content of some affective impression, free from the grip of fate, since

273 those old philosophers who used to say that everything takes place by fate held that assent is given by force and necessity. But those who disagreed with them released assent from fate and denied that if assent were tied to fate it would be possible to disentangle it from necessity. They argued as follows: (*a*) if all things occur by fate, all things occur by an antecedent cause; and (*b*) if desire is caused, those things which follow desire are also caused; therefore (*c*) assent is also caused. But (*d*) if the cause of the desire is not situated within us, even desire itself is not in our power; and (*e*) if this is so, those things which are caused by desire do not rest with us. (*f*) Thus it follows that neither assent nor action are in our power. Hence (*g*) there is no justice in either praise or blame, honours or punishments. (ibid. 40)

A version of this argument was considered in the previous chapter (VI.3*a*, **233**), and it amounts to an expanded version of the inference from determinism to human unfreedom (and contrapositively from freedom to indeterminism) noted above. Chrysippus tries to evade either result by questioning the argument's validity, a manœuvre he attempts precisely by distinguishing antecedent from internal causes (§§1*d–f* above).

He is happy to allow that our desires are antecedently caused, as are our assents to them: ‘assent cannot take place unless prompted by a sense- impression’ (ibid. 42). But antecedent causes do not determine what

results from them on their own, but only in conjunction with the internal states and conditions they operate upon. Our assents are thus (partially) caused by our impressions, but they are not determined by them (since we assent or fail to assent on account of our dispositions). Chrysippus, then, accepts (*c*); but on the proper analysis of the cause of assent, it is not the case that ‘the cause of the desire is not situated within us’; it is the fact that we are constituted as we are that makes us assent to a particular impression when it arises. Thus (*d*) may also be true, but since its antecedent is false, the argument fails.

However, this causes immediate difficulties. While Chrysippus' identifi-cation of fate with antecedent causes is well attested (ps.-Plutarch, *On Fate* 11. 574d; cf. Cicero, *Topics* 59), Plutarch (*On Stoic Self-Contradictions* 47. 1056*b–c*) protests that if Chrysippus makes fate merely the antecedent (rather than the perfect) cause of right action, his position will become untenable, since antecedent causes are supposedly weaker than perfect causes, yet nothing is more powerful than the will of Zeus (which Chrysippus also equates with fate); moreover, fate is, as we have seen, supposed to be ineluctable and unconquerable.

There are a number of ways to try to resolve this.[34] It has been suggested that Chrysippus is attempting to show that assent is not necessitated; but as even this interpretation's principal partisan[35] notes, it is hard to see how the argument for that might go. Alternatively, Chrysippus may rather be denying that the fact of the (partial) antecedent causation of our assents removes any moral responsibility we may have for them (since our dispositions, which are not external to us, play a crucial role);[36] and no doubt Chrysippus does wish to uphold this. But to do so, he must also show that either we are responsible, in a strong sense, for our dispositions (as Aristotle tried to do: IV.4*b*), or that the crucial relation of responsibility, unlike ordinary causality, is not transitive (i.e. that even if something external to us is responsible for our characters, *we* are still fully responsible for how we act according to them): and that is no mean feat either.

Most probably, however, Chrysippus simply wishes to deny that fate in and of itself necessitates our actions. Given the structure of our dispositions, fate is a necessitator. In general, given objects' natural potentialities for behaviour, along with sundry contingent facts about their physical and spatial relationships (contingent in the sense that they are not species-necessary for the objects in question), a Laplacean super- scientist will be able to predict all their future interactions, since those futures are determined by the way in which each reacts to the others, the intersections of their spatio-temporal trajectories being the antecedent

causes which trigger their respective dispositions to act (this picture is at least compatible with Alexander's account at *On Fate* 22. 191. 30–193. 2: see X.4*b*).

Even so, things might (in one sense) have turned out otherwise: indeed they would have if the initial conditions had been different. Of course, causally speaking those conditions could not have been other than they were (the will of Zeus really is ineluctable); but we can still assess what would have happened had they been. Had it been you rather than me who was tempted with the third helping of pasta, you would have refused, since you value your figure more than immediate gustatory gratification, whereas I do not. Nor is there any difficulty about the truth-conditions for these counterfactuals: we need only consider previous real-world situations similar in all the crucial respects (in this case, previous instances of third-helping offering). Possibility need not collapse into necessity, and nor must it be merely epistemic: rather it derives from the actual physical propensities of things.

(c) Fate and Responsibility: Confatalia

However, it is far from clear whether such a picture will allow Chrysippus to maintain that we are, in any strong sense, responsible for what we do. Even if we allow that the distinction between internal and external causes, and the priority of the former over the latter, shows that the things we do are ‘up to us’ in some sense (since our particular beliefs and desires causally determine our particular reactions to stimuli), it may not be enough to underwrite our institutions of praising and blaming, since the formation of the crucial desires and beliefs may itself be outside our control, as Aristotle realized (IV.4*b*). Our actions may still be attributable to us, just in that it is we who perform them, which is enough to rebut the Epicurean charge that this kind of determinism can make no sense of the notion of coercion (although the subtler point that there will be no *morally relevant* difference between coercion and voluntary action may still be defensible). But that Humean response does not explain how we may justify the ascription of responsibility to an individual, given that everything is necessitated by fate, even if an individual's actual dispositions are part of the causal nexus that determines those outcomes.

The Stoics held us to be responsible for our actions because our natures determined our choices of action.[37] But that invites an obvious rejoinder:

274 if they assign impulse as being up to us on the grounds that we have it by nature, what is to stop us from saying that burning is up to fire since fire burns by nature? (Nemesius, *On the Nature of Man* 35 = *SVF* 2. 991)

The Stoics apparently thought that the more complex the internal causal process linking impression and action, the more attributable to the individual it becomes. But there is no obvious reason (phenomenology apart) for thinking that any such increase in complexity will generate the emergence of responsibility in the requisite sense. My computer is far more complex than my calculator; and it is, in a sense, responsible for performing correctly the tasks I assign it. But if it fails to do so, I am no more justified in holding it culpable than I am my calculator (even if, in a fit of Basil Fawlty-ish anthropomorphism, I berate and belabour it): the blame, if any, attaches either to me for programming it wrongly, or to its manufacturer.

Perhaps the Stoics need to abandon responsibility and retreat into hard determinism.[38] The famous story about Zeno, that

275 he was once beating a slave for stealing, and when the latter said 'I was fated to steal', he replied 'and to be flogged', (Diogenes Laertius, *Lives* 7. 23 = *SVF* 1. 298 = 62e LS)

is compatible with hard determinism, but it does not entail it. In fact, it is rather intended to illustrate the Stoic doctrine of 'confatalia' (as Cicero calls them), or 'co-fated events', a doctrine designed to counter the so- called 'lazy argument':

276 they pose it thus: 'if it is your fate to recover from illness, you will recover, regardless of whether or not you call the doctor. Similarly if it is your fate not to recover . . . you will not recover whether or not you call the doctor. And one or the other *is* your fate. Therefore it is pointless to call the doctor.' (Cicero, *On Fate* 28–9 = 55s LS)

That argument mistakenly assimilates determinism to fatalism, the idea that a particular fate will befall you no matter what you do to evade it. Chrysippus replied that

277 some events in the world are simple, some complex. 'Socrates will die on such and such a day' is simple: his day of dying is fixed, regardless of what he may or may not do. But if a fate is of the form 'Oedipus will be born to Laius', it will not be possible to add 'regardless of whether or not Laius has intercourse with a woman'. For the event is complex and 'co-fated'. He [i.e. Chrysippus] uses this term because what is fated is both that Laius will have intercourse with his wife and that by her he will beget an Oedipus. . . . All fallacies of this sort are refuted in the same way. 'You will recover regardless of whether or

end p.258

not you call the doctor' is fallacious. For it is just as much fated for you to call the doctor as for you to recover. (ibid. 30 = 55s LS)

In general, on this view, if some event is fated to occur, then so too are all its necessary conditions. I accept the third helping because I am greedy, a condition which for me is fated just as much as my eating it. Moreover, my gluttony explains my eating it. But the simple fact that in one (causal) sense I could do no other does not empty my actions of their point. Even in a totally determined universe, persistent backsliders can still try to mend their ways; and they may even succeed, through their own efforts, in doing so (although their making the effort is itself determined). Whether or not they do so will be determined by causes remote from their own control; but we may still ascribe to them both their acratic condition and their determination to do something about it (see further XI.5).

(d) Divination and Fate

The questions of whether the universe is deterministic and whether its future can be known are logically distinct (VI.6*a*). But in general the ancients agreed that for some future outcome to be known there must be determinate causes of it presently existent and knowable; and on that assumption, accurate and general forecasting of the future will require that the future be determined. By and large the Stoics believed in divination (although Panaetius was an exception: Cicero, *On Divination* 1. 6, 2. 88, 97), which they thought mutually supportive of their commitment to determinism, a position some thought question-begging:

278 Chrysippus gives this demonstration to us, proving each one by way of the other. For he wants to establish that everything comes to be according to fate on the grounds of divination, while that divination exists he is able to show by no other means than by assuming that everything comes about according to fate. (Diogenianus, in Eusebius, *Preparation for the Gospel* 4. 3 = *SVF* 2. 939 = 55p LS)

But there is no vicious circularity here: divinatory successes provide empirical support for the hypothesis of causal determinism, which in its turn accounts for the possibility of divination (cf. XI.4*d*).

Even if the universe is deterministic, its future may still not be humanly predictable, since the number of variables involved in any such prediction may be beyond human computability. But it may still be accessible to divine knowledge, and on the assumptions that such knowledge would be beneficial for human beings and that the gods care for human affairs, we might expect that the gods would find other ways of letting us know how things are going to go (cf. Cicero, *On Divination* 101–2; *On the Nature of the Gods* 2. 161–8). Those assumptions, both explicitly endorsed by the

Stoics, are both questionable, and were attacked by the Stoics' opponents (cf. Carneades, in Cicero, *On the Nature of the Gods* 1. 4; Favorinus, in Gellius, *Attic Nights* 14. 1–36).[39] But Stoic theology and metaphysics are at least mutually consistent, supportive even. The debate on divination will be treated more fully in the following chapter (VIII.5*a*–*b*; cf. XI.1*b*; XII.2*d*); but one feature of it deserves consideration here. The Stoic definition of divination was ‘prophecy of events that occur by chance’ (Cicero, *On Divination* 2. 13–15, 26); yet there seems no place for chance in the Stoic universe. Indeed the Peripatetic attack on the Stoics' determinism takes particular aim at their refusal to allow the existence of genuine chance (Alexander, *On Fate* 8; *Mantissa* 24; see X.4*b*–*c*). Cicero accuses them of self-contradiction: only if something is going to occur can we predict it; but if it is going to occur, it is now impossible that it fail to occur; hence if predictable, it cannot be a matter of chance. And the only basis for any such knowledge is causal: the event is deducible from known initial conditions by way of causal laws. But then too it cannot be chance.

The Stoics have an easy rejoinder to this. They define chance epistemically as 'a cause obscure to human understanding' (§2*a* above). An event is chance just in case we cannot predict it on the basis of known laws and conditions. It may for all that be causally predictable by a divine super-intelligence, who may also offer us signs of its future occurrence. But these signs are not supposed, for the Stoics, to be indicative signs: that is, no direct causal ties are presupposed between the sign and what it signifies, nor is the latter rendered necessary in any strong sense by the former (on sign-theory, see VI.4). Consider the 'astrological theorem'

(T) if someone is born at the rising of the Dog-star, he will not die at sea.

If that is a universal truth, then

(I) if Fabius was born at the rising of the Dog-star, then Fabius will not die at sea

ought, as a substitution instance of T, to be a sound conditional. But, on the Stoic account of the truth-conditions of the conditional,[40] antecedent and consequent must be connected by a close tie of relevance: the antecedent must make it the case that the consequent is true. Yet whatever else may be true, Fabius' particular birth-aspect hardly seems to be a cause of his dodging death by drowning. For this reason Chrysippus held that such propositions should not be formulated as conditionals at all, but rather as negated conjunctions:

end p.260

(I*) it is not the case both that Fabius is born at the rising of the Dog-star and that he dies at sea.

Cicero pokes fun at this (*On Fate* 15); what is to prevent doctors and geometricians from doing the same:

279 what cannot be carried over in that sort of way from the form of a necessary consequence to that of a negated conjunction? (ibid. 16)

But Cicero fails to see that propositions such as (1*) (by contrast with those like (1)), suggest no direct and necessary connection between antecedent and consequent. They may, for all that, be causally related, e.g. as collateral effects of some further cause (and indeed of the general *sumpatheia* the cosmos exhibits: §1*b* above; compare the issue between 'soft' and 'hard' astrology: VIII.5*b*; XI.1*b*; XII.2*d*). *Pace* Cicero, Chrysippus' claim that while (1*) is true (1) is false is serious and pointed.

(e) Soft Determinism

The Stoics had no desire to exonerate the wrongdoer by way of their determinism, and their austere morality emphasizes the importance of individual striving towards self-improvement and moral progress. But this stance is not without its problems. The determinist can attempt a causal explanation of the occurrence of such institutions as praising and blaming, and may (in evolutionary terms) give an account of their beneficial effects. Deterrence as justification for punishment, *pour encourager les autres*, is at least indifferent towards determinism (it may even favour it). Nor does determinism render pointless attempts at criminal rehabilitation, or social protection. Even so, in spite of Chrysippus' efforts to drive a wedge between hard determinism and Epicurean causal chaos, the deeply rooted idea that people *deserve* to be treated in a certain way according to how they behave seems unreconcilable with a deterministic universe.

The Stoics seek to defuse that difficulty in a variety of ways. They argue that moral judgements are themselves necessary products of the natural order, and so they too must be natural. Praise and blame are part of human nature. Man is a censorious animal:

280 they [the Stoics] suppose that everything naturally constituted is such as it is in accordance with fate, 'natural' being the same thing as 'in accordance with fate', and they add 'consequently it will be in accordance with fate that animals have perceptions and impulses. And some animals will be merely active, while others will perform rational actions. And some will do wrong, while others will do right actions. For these are natural to them. But so long as wrong and right actions remain, and their natures and qualities are not

end p.261

removed, there also remain commendations and censures, punishments and honours. For such are the sequence and order to which they are subject.' (Alexander, *On Fate* 34. 205. 24–206. 2 = *SVF* 2. 1002 = 62i LS; cf. ibid. 35. 207. 5–21 = *SVF* 2. 1003 = 62j LS)

That people will perform some right and some wrong actions is no less causally determined than anything else, as too are the consequences of those actions. Zeno's slave (**275**) has no grounds for complaint. Even so, as Alexander goes on to ask, it is difficult to see how the concepts of right and wrong can retain their content in such a universe. Actions are morally appraisable only if they are up to us in the sense of not being compelled by events outside our control. But whatever the Stoics say, in their world all actions are forced in this way (ibid. 34. 206. 2–207. 3). It is not enough, Alexander thinks, merely to show how such institutions might be causally explained, since such explanations do not account for their essential moral value. Indeed such genetic explanations undermine it: for they amount to providing a natural history for the concepts in question, and hence, as far as their content is concerned, to explaining them away. There is no genuine alternative to hard determinism.

Alexander is right that some slippage in the content of the moral notions is unavoidable for the Stoics. But it is not yet clear whether that shift need be so radical as to drain them of all genuine content. In order to see how damaging it, is we need to sketch the Stoic account of the overall structure of the cosmos.

3. Stoic Cosmology

(a) The Rationality of the World

The Stoics' universe is thoroughly materialist in character. Even the intelligence which permeates it does so in the form of a particularly fine substance, the *pneuma* whose role as a containing cause of material objects has been considered earlier (§ 1. *b*). Indeed, they identify this intelligence with god:

281 he [Chrysippus] says that divine power resides in reason and in the mind and in the intellect of universal nature. He says that god is the world itself and the universal pervasiveness of its mind; also that he is the world's own governing faculty, since he is located in intellect and reason; that he is the common nature of all things, universal and all-embracing; also the force of fate and the necessity of future events. Moreover, he is fire, and the ether of which I spoke earlier, and things in a natural state of flux and mobility, like water, earth, air, sun, moon, and stars, and the all-embracing whole. (Cicero, *Nature of the Gods* 1. 39 = *SVF* 2. 1077 = 54b LS)

end p.262

The Heraclitean overtones of **281** are clear enough (cf. I.3*b*),[41] although the pervasiveness of the divine mind also recalls Anaxagoras (I.4*b*). Moreover, the Stoics are four-element theorists. But more important than the debts they owe to their predecessors is their extraordinarily powerful, indeed positively Panglossian, directed teleology:

282 there is nothing apart from the world which lacks nothing, and which is equipped from every viewpoint, perfect, and complete in all its measures and parts. As Chrysippus cleverly put it, just as the shield-cover was made for the sake of the shield and the sheath for the sake of the sword, so too with the exception of the world itself everything else was made for the sake of other things; for example the crops and fruits brought forth by the earth were made for the sake of animals, and the animals for the sake of men (the horse for transport, the ox for ploughing, the dog for hunting and guarding). Man himself came to be in order to contemplate and imitate the world, being by no means perfect, but a tiny part of that which is perfect. (ibid. 2. 37 = 54h LS)

Some of this recalls Aristotle, particularly in his more anthropocentric moments (IV.3*d*); but the god of **281** is thoroughly un-Aristotelian. Aristotle too held that Nature was an internal power within things; but for the Stoics that power is divine, and each individual nature is part of Nature as a whole in a very strong sense. In stark contrast with Aristotle (although in line with Plato of the *Timaeus*) they conceive the world as a single living organism, all of whose parts contribute to the overall functioning of the whole, and all of which is driven by the perfectly good divine reason.

Thus Chrysippus held notoriously that 'bedbugs are useful for waking us, and mice for encouraging us not to be untidy' (Plutarch, *On Stoic Self- Contradictions* 1044d) while the function of the pig was to be eaten (Porphyry, *On Abstinence* 3. 20. 1). The Stoics' principal target is Epicurean pure mechanism (cf. VI.2*d*, **226**), since the evident construction of the world shows it to have been the work of a divine artificer:

283 suppose someone were to bring to Scythia or Britain the armillary sphere recently built by our friend Posidonius which . . . brings about in the sun, the moon, and the five planets effects identical to those brought about day by day and night by night in the heavens. Who in those foreign lands would doubt that the sphere was a product of reason? And yet these people hesitate as to whether the world, from which all things come into being, is itself the product of some kind of accident or necessity or of a divine mind's reason. And they rate Archimedes' achievement in imitating the revolutions of the heavenly sphere higher than nature's in creating them—and that when the original is a vastly more brilliant creation than the copy. (Cicero, *On the Nature of the Gods* 2.88 = 54i LS)

end p.263

But if it has been constructed, it must have been made for something's benefit, in this case ‘for those animate creatures which use reason: i.e. gods and men’ (ibid. 133; cf. X.2*d*; XI.3*b*).

At this point, evil becomes a problem. If the world is the supremely good result of divine design, why are some of its contents less than good? Chrysippus answered that good could not exist without evil, since the two are correlatives (Aulus Gellius, *Attic Nights* 7. 1. 1–7), a view which has had many adherents; but I can see nothing to be said in its favour (it is one thing to claim that we could have no concept of good unless there were some evil in the world, quite another to hold that there could *be* no good under those circumstances). Rather more substantially, Chrysippus makes moves which parallel Plato's in the *Timaeus* (III.4*a–b*): if humans are to be made rational, they must be constructed out of thin material, which is naturally susceptible of disease and decay (ibid. 8–13). And some apparent evils (like bedbugs) are in fact goods in disguise.

Such defences did not command widespread support. But even more shocking was felt to be the Stoics' view that everything in the world was for the best, including, by consequence, the existence of vice (Plutarch, *On Stoic Self-Contradictions* 1050*c–d*). In order to answer the questions posed at the end of the last section the Stoics must rely heavily on two quite distinct notions: that individual creatures are as good as *they* can be; and that the world, and all its contents, is in a constant state of evolution towards a better end.

The Stoics borrow from Aristotle the notion of natural hierarchy of functions; and like him they interpret this in a teleological fashion. Different animals do different things: and what they do uniquely or best is their proper or definitive function. It is their nature for them to act thus; and hence it is right for them to do so. Man censures; hence it is right for man to be censorious. But the notion of rightness involved here seems ambiguous. Perhaps ‘right’ just means ‘fitting’ or ‘appropriate’; thus it is right for a carving-knife to slice meat, and not right for it to prune vines, or swat flies. And you might, at a pinch, reproach someone for misusing it in one of these ways. But no one would censure the knife for being so misused, since it is merely instrumental; yet it is hard to see how human beings, in the Stoic scheme of things, are any less instrumental, or have any morally richer role to play.

On the other hand, the Stoic universe is not merely a hierarchy of functions, it is a hierarchy of *good* ones, underwritten by divine providence. But then we may view the world's steady progress towards perfection as being itself a good thing, and there will be a further sense distinct from their simple functional fitness in which it is good that knives cut, or men blame, since these actions are part of the providential order of things. It is no mere fact that we praise and blame; rather praise and blame are part of the causal structure of the best of all possible worlds, and an

integral component in its struggle for improvement. But how then may the Stoics distinguish the good from the bad? If the world is the best it could be (cf. Epictetus, *Dissertations* 1. 1. 7–12, following Plato, *Timaeus* 29e–30b, 75a–c) and is evolving towards perfection, then, given that everything which occurs contributes towards that evolution, every event must be good. And indeed the Stoics accepted that (Marcus Aurelius 4. 10, 23, 26; cf. 2. 3).

Even so, they may yet develop a non-trivial account of the way in which particular actions and events may be relatively good or bad, by distinguishing between temporally indexed and timeless judgements of worth. Consider a paradigm case of an action that would be considered wrong, by the Stoics as well as everybody else (Stoic morality was notoriously unconventional: Sextus Empiricus, *Outlines of Pyrrhonism* 3. 200–1, 205–6, etc.). Suppose *a* murders *b* at spatio-temporal region *s*. The action may be considered either as (*a*) a murder *tout court*, or as (*b*) a murder- at-*s*. Other things being equal, murder is wrong, and so *sub* (*a*) we condemn it; however viewed *sub* (*b*), it is an event conducive towards, because part of the causal nexus leading to, the eventual perfection of the world; thus we may welcome it, not because it *is* a murder (murders are still reprehensible), but simply because it is so conducive. We may flesh that out a little: what makes murder (*sub* (*a*)) wrong is the fact that a perfect world will contain no murders, even if *sub* (*b*) it is a good thing that it happened.

Thus the Stoics may salvage a non-trivial sense in which actions can be genuinely evaluated. Yet it is one thing to assess the act as such, and quite another to hold the agent responsible for it. Here the Stoics' opponents are on firmer ground. As a Stoic, I may think you a frightful bore; I may wish that you weren't; and I may anticipate with satisfaction the blessed day when the world contains only interesting people. But it still seems unfair to hold *you* any more responsible for the tediousness of your company than for the fact that you are hideous, stunted, or deformed.

(b) Fate and Moral Progress

Stoic determinism is, via the doctrine of confatalia, distinct from fatalism, the idea that your fate is fixed no matter what you do to evade it. None the less, the Stoics sometimes stray dangerously close to fatalism: Zeno and Chrysippus

284 said that everything is fated, using the following example. If a dog is tied to a wagon, if it wants to follow, it is both pulled and follows, making its voluntary action coincide with necessity; but even if it does not want to follow, it will be forced to do so anyway. So too with human beings: even if they do not want to, they will be forced to follow what is destined anyway. (Hippolytus, *Refutation* 1. 21 = *SVF* 2. 975 = 62a LS)

end p.265

According to **284**, the Stoics think that human beings, while unable to determine how things will be, can none the less choose, in some relatively strong sense, whether to accept them. Our choices are thus real, but causally insulated from the way the world works. But that picture sits badly with our earlier analysis of the Stoic structure of events. The illustration offered in **284** does have unfortunate fatalistic overtones, but it may still be interpreted consistently with a non-fatalist account of causation.

Its principal point concerns our stance in regard to our goals. The Stoics held that, until we achieve the condition of cognitive perfection of the sage who never makes any mistakes, every time we express a desire to do something we should mentally qualify our project with what they called a ‘reservation’ (*huphexairesis*). We should want something only if it accords with the will of Zeus (and hence will occur as part of the fated unravelling of the universe). Reservation amounts to qualifying all expressions of desire with a mental ‘God willing’; and, crucially, that reservation is supposed to be part of the actual desire itself. We should no longer simply desire something: we should desire it only on condition that its realization is part of the structure of fate. The goal of Stoic moral and psychological development is to make one's own impulses and desires harmonize as closely as possible with the way things are actually going to turn out. The Stoic will have no unconditional desires, and hence will have no real desires for what is (as a matter of causal fact) impossible.

Such a position will indeed be fatalistic, on the assumption that the way things will turn out is determined quite independently of human choices and decisions. But there is no reason to make any such assumption. On the contrary, those choices and decisions are part of the instrumental causal structure by which fate, the will of Zeus, the overall *logos* of things, is worked out and brought to fruition (compare Cleanthes' surviving *Hymn to Zeus*: *SVF* 1. 537 = 54i LS). In the same way that murders, although in themselves bad, can be conducive to the overall good, so my frustrated desires, in their very frustration, may contribute to the eventual perfection of the universe. Even so, the closer I (and the world) approach moral perfection, the less my desires will as a matter of fact be frustrated. That position is both coherent, and non-fatalistic, since it does not involve the counterfactual claim that things would have turned out the same whatever my decisions may have been. It is, of course, a further question whether there is anything else to commend it.

4. Conclusions

The Stoics insist that causation is a physical relation between bodies; and for this reason they are committed to a thoroughgoing materialism: everything
end p.266

in the world, including God and the mind, are material (§§1*a*–*b*). They produce an elaborate scheme of causal classification, designed to explicate their determinism as well as to allow for the temporal sequentiality of causation (§§1*c*–*g*). This determinism is developed in such a way as to allow for genuine possibility (§2*a*); all events are necessitated, but they are not all (from the point of view of the substances which instantiate them) necessary (§2*b*). None the less, the Stoics seek within this framework to hold humans accountable for their actions: what we do is up to us, in the sense that it is the product of our beliefs and dispositions—we are not simply pawns on the cosmic chessboard (§§2*c*–*e*). The Stoics' belief in the fundamental rationality of the universe (a result of its permeation by the vital, intelligent *pneuma*, which is, as a whole, God) makes them subscribers to a powerful form of directed teleology: the world is created, and is the best possible (§3*a*), a belief they seek to reconcile both with the claim that it embodies progress, and with the claim that (from a certain point of view at any rate) some features of it are, objectively speaking, bad (§ *b*). They thus embrace determinism with more commitment than any of their predecessors; and they unflinchingly and rigorously deduce, and accept, the consequences of it.
end p.267

VIII The Sceptics

R. J. Hankinson

1. The Development of Scepticism

Sceptical reflections on the scope of human knowledge begin with Xenophanes (I.3*a*); but there is little evidence of any systematic scepticism of the senses until Aristotle somewhat intemperately dismisses it in his *Metaphysics* (4. 4–5). Aristotle's targets are unnamed, although at roughly the same time Pyrrho of Elis (*c*.360–*c*.270 bc) propounded a sceptical philosophy of life; it is said (no doubt apocryphally) that he was saved from walking over precipices and under oncoming traffic only by the solicitude of his friends (Diogenes Laertius, *Lives* 9. 62). Pyrrho's concerns were, however, practical; he seems to have had little interest in combating the theoretical pretensions of the Dogmatists, as later Sceptics were dismissively to style their philosophical and Scientific opponents.[1]

In 272 bc Arcesilaus (*c*.318–*c*.243 bc) succeeded to the headship of Plato's Academy. Reacting against the increasingly sclerotic nature of school Platonism, he sought to re-establish its doctrinal purity by emphasizing its Socratic origins in the practices of questioning and refutation. Under his direction the Academy was devoted to destructive dialectic and the exposure of incoherence, singling out for particular attack the Stoics' optimistic epistemology, with its doctrine that some perceptions (the ‘cataleptic’

impressions) were self-guaranteeingly true.[2] But they apparently spent little time directly attacking the various Dogmatic accounts of cause and explanation. It is not until Aenesidemus (fl. *c*.75 bc) abandons an Academy by now sceptical in little more than name in order to refound Pyrrhonism that we find evidence of a systematic attempt to undermine the pretensions of the Dogmatists to explanatory success.[3]
end p.268

2. Aenesidemus Against the Aetiologists

None of Aenesidemus' work survives, although we possess a summary of his *Pyrrhonian Discourses* in the ninth-century Byzantine patriarch Photius' library catalogue. Photius writes that Aenesidemus

285 in his fifth volume holds out an aporetic guard against causes, refusing to concede that anything is the cause of anything, saying that the aetiologists are mistaken and enumerating some modes according to which he thinks that, by being attracted to causal theory, they have been steered into such an error. (Photius, *Bibliotheca* 212. $170^b17–22$ = 72l LS

'Modes' or 'tropes' (*tropoi*) are general patterns of argument applicable variously to particular cases; and Sextus Empiricus (*Outlines of Pyrrhonism* 1. 180–5) ascribes to Aenesidemus Eight Modes 'against the aetiologists', or causal theorists. Aetiology is the supplying of causal explanations, *aitiai*, for things, typically involving hidden causes, deep reasons, and subtle explanations. Aetiology was particularly associated with the Peripatetics; Strabo says of the first-century bc eclectic Stoic Posidonius that:

286 there is in him an abundance of aetiologizing and Aristotelizing, which our school avoids because of the obscurity of the causes. (Strabo 2. 3. 8 = T85 EK;[4] see further X.2*a*)

And while the practice of giving *aitiai* was by no means restricted to philosophers and scientists (the third-century bc Alexandrian poet Callimachus produced a collection of *aitiai*, or mythological explanations), it is against them that Aenesidemus trains his guns. Sextus is our only substantial source for the Eight Modes,[5] and his account is brief and compressed. But it is evident even from his summary that Aenesidemus developed a general attack on philosophico-scientific pretensions to explanation of great power and originality. Some of his terminology suggests that the Epicureans were his main target; but the arguments work indifferently against any unduly optimistic scientific epistemology that seeks to penetrate beneath the surface of things to give an account of their basic structures.

Let us examine the Modes, more or less in order:

287 the First is the mode according to which, he says, aetiology in general, being concerned with non-apparent things, has no agreed confirmation (*epimarturēsis*) from the appearances. (*Outlines of Pyrrhonism* 1. 181 = 72m LS)

end p.269

The First Mode raises difficulties for any inference from appearances to hidden states of affairs, of a sort attempted in indicative sign-inferences (VI.4*a*). Since the hidden reality to which we infer is by hypothesis unobservable, we have no independent means of testing the validity of our inferences, since we cannot actually uncover the mysterious hidden world. This problem is a serious one on the assumption that there is more than one consistent 'explanation' for any particular set of phenomena, an assumption that was rejected in some cases. The Atomists thought atomism to be the only hypothesis capable of saving the appearances of motion and plurality from Eleatic assault (VI.1*c*, 2*c*), while the Stoic inference from the phenomenon of sweating to the existence of invisible cutaneal pores was a paradigm indicative sign (VI.4*a*). If only one explanation is in fact consistent in this way with the observable phenomena, Dogmatists are indeed entitled to claim that the appearances entail their particular picture of the world's basic structure. But the prospects for a successful quasi-transcendental argument of this sort from matters of empirical fact to the only conceivable explanation are less rosy than sanguine Dogmatists pretend:

288 the Second Mode shows that frequently when there is an abundance of ways of assigning an explanation to what is under investigation, some of them account for it in one way only. (Sextus Empiricus, *Outlines of Pyrrhonism* 1. 181 = 72m LS

Dogmatists may pretend that theirs is the only explanation available, and that it is thus entailed by the phenomena; but they are deluding both themselves and others. Even though it is clear that many different and incompatible explanations can frequently be given for the same phenomenon, Dogmatists (of whatever stripe) resolutely ignore the competing accounts in favour of their own (a partial exception is the Epicurean doctrine of multiple explanations: VI.2*f*). And that is no doubt at least sometimes true.

The idea that theoreticians are in perpetual and undecidable disagreement is fundamental to the Pyrrhonist enterprise: it is because, Pyrrhonists say, that disputes (*diaphōniai*) are everywhere endemic among philosophers and scientists, and that there can be no principled and non-question- begging way of settling them, that they are forced to suspend judgement on every 'non-phenomenal question of scientific inquiry' (ibid. 1. 13). Indeed an appeal to the ubiquitousness of dispute (or at least potential dispute)[6] forms one of another set of canonical sceptical modes, the Mode

of Dispute from the Five Modes of Agrippa (§4 below). So the first Two Modes operate in tandem. The first asserts that one cannot simply peel off the veil of the phenomena to examine the reality underneath. Phenomena, being the type of things that they are, cannot have any tendency even to render plausible any conjectures about the status of a completely different category of things, the real underlying objects (this is Berkeleian in tone: *Principles* §§8, 18–20); while the potential proliferation of competing theories adduced by the Second Mode rules out the possibility of a single one simply winning by default.

There are a number of things Dogmatists might say in reply to that. First of all, they may urge, even if the underlying reality now lies hidden, more advanced technology may in future allow us to strengthen and amplify our perceptual information to make it patent. Consider the case of sweating and the pores. For the ancients it was necessary to infer the existence of the latter by way of the former (invoking argument [A]: VI.4*a*), since the pores themselves lay below the threshold of observability. But the microscope has made them visible to us, making new appearances available to confirm the inference. Equally Galileo's telescope provided hard observational evidence (of the moons of Jupiter and the phases of Venus) to disconfirm the traditional Ptolemaic–Aristotelian world-picture and hence to corroborate the Copernico-Keplerian model of the solar system.

But even if that story seems eminently reasonable to us, it need not have done to the ancients: for they had no experience of perception-enhancing instruments, and consequently no reason to set any store by their possibility. In any case, there is a difficulty surrounding the whole notion of enhanced perception: how can we be sure that the instruments do not distort instead of enhance reality? Galileo himself had great difficulty persuading people that his telescope was in fact a more or less accurate means of intensifying visual perception since there was no plausible optical theory available at the time which could account for such effects. Moreover, the images of Galileo's telescopes were crude in the extreme, disturbed by ghost-images and diffraction-lines caused by the extremely primitive state of lens-grinding. And as soon as theory is invoked to explain some effects as veridical, others as mere interference, it is itself open to challenge.

The ancients were not unaware of the difficulties under which they laboured. Galen, for one, was acutely conscious of his lack of precise instrumentation. He believed that the senses could be trained in discernment; but there were limits to the practicability of this procedure, and where it gave out Galen thought that the raw evidence of the senses had to

end p.271

be supplemented by inference. In his *Anatomical Procedures*[7] (9. 7. 6 Simon), Galen ‘proves' the existence of invisible pores in the surface of the meninges on the grounds that when held up to the light they are translucent. But that inference would not satisfy any sceptic, and rightly so: it rests upon further assumptions regarding the nature of the propagation of light which are (at least *grosso modo*) false; and so, consequently, is the inference.

On a rather different tack, the Dogmatist might argue that, even if sensory experience cannot ever directly confirm or disconfirm theories about the subsensory world, we may yet infer from the theory's predictive success that it is very likely roughly right. Verisimilitude is a notoriously problematic notion;[8] but even if it can be made coherent, the sceptic here will reply, in the manner of the Empiricist doctor, that *all* that matters here is success: and that can be gained without any theoretical substructure. All we need are empirically supported connections between antecedent and subsequent events; and those events will be, in their very nature, phenomenal. The detour through theory is simply a waste of time, and encourages the fantasy that we can know things we cannot. The appearance of the Epicurean technical term *epimarturēsis*, or confirmation, in **287** suggests that Aenesidemus' attack was directed particularly against the Atomists. Epicurean *epimarturēsis* (see VI.4*b)* consists in corroborating weak or equivocal sense-reports on the basis of more unambiguous and secure data. In the example deployed (to different ends of course) by both Sceptic and Epicurean, if a square tower can appear round from a distance, and from a distance I see one which looks round, I must approach more closely to determine whether it really is round or square. Even in ordinary cases of perception, the Sceptics challenge the validity of such a procedure; but as far as fundamental, theoretical entities are concerned, matters are clearly far worse: no refinement of the perceptual circumstances can offer us any confirmation whatever.

It is, however, unclear whether the Epicureans intended the theory of *epimarturēsis* to apply to cases of confirmation of the theoretical substructure. Most of the surviving relevant texts confine its scope to the confirmation of things at least potentially evident to the senses by way of other empirical evidence. **287** extends the concept beyond its usual confines within the ambit of commemorative signs (VI.4*a*). It is not easy to determine whether this extension is legitimately Epicurean, or simply Aenesidemus' invention.[9] But even if it is, the Epicureans are still in genuine

difficulties. If they do expand the range of *epimarturēsis* to include the confirmation of intrinsically non-evident objects, they draw the fire of the First Mode; if they do not, on the other hand, it is hard to see how they can go about supporting the plausibility of their views about the natures of things, given the problem of underdetermination raised by the Second. Whatever the nature of the Atomists' views, the first two Modes represent a significant challenge to any realist theory of fundamental entities and structures.

289 The Third [Mode] is that according to which they assign to orderly comings- to-be (*ginomenōn*) causes exhibiting no order' (Sextus Empiricus, *Outlines of Pyrrhonism*. 1. 182).

Sextus says no more about the Third Mode. Bury (1933: 105) translates *ginomenōn* simply as 'events', a sense it can perfectly well bear in Greek. But Aenesidemus' principal concern here is probably with the explanation of generation. As we have seen, Atomists (indeed mechanists in general) had enormous difficulty in explaining the continuity of natural physical processes, paradigmatically animal reproduction, on the basis of a fundamental physics of random atomic interaction (VI.2*d*; VII.3*a*). In the case of hugely complex yet precisely repetitive processes, the mere assignation of some particular atomic cause to each individual event that makes up the process will be explanatorily inadequate, since it is precisely the regularity of the whole structure that requires explanation.

The point is generalizable to cover all types of processes, and need not be confined to generation as such; and it recalls Aristotle's discussion of coincidence (IV.2*b*). Even if there is an explanation for *p* and an explanation for *q*, it does not follow that their conjunction explains the conjunction of *p* and *q*. I came to the market to buy grain, you to sell pigs—but that does not explain our being there at the same time so I could collect the money you owed me (*Physics* 2. 5. 196^{b}35–197^{a}5; cf. *Metaphysics* 6. 3; IV.2*b*). That lack of an explanation may be all very well in the case of genuine coincidences: that is just what a coincidence is. But nature as a whole cannot simply be a matter of chance in this way.

This is a serious charge. In many cases mechanistic science of the Epicurean (indeed of the Empedoclean) kind could 'explain' only the parts, leaving the orderliness of the whole quite unaccounted for (VI.1*d*), which inadequacy (basically the inability to explain the emergence of form on the basis of an inherently formless substructure) gave the impetus to ancient teleology in all its forms (III.1*a*, 4*a*; IV.3*a*; VII.3*a*; X.3*a*).

290 The Fourth [Mode] is that according to which, having seen how the appearances come to be, they imagine that they have also got a grip on the way things non-apparent come to be; and while perhaps the non-apparent are brought about in the same way as what is apparent, perhaps on the other hand they are

end p.273

not, but come to be in their own peculiar fashion. (Sextus Empiricus, *Outlines of Pyrrhonism* 1. 182 = 72m LS)

Again we need to flesh out Sextus' compressed report. **290** might be interpreted as taking issue with the sort of extremely naïve realism which postulates an equivalence between observable, phenomenal properties and the real properties of things. Veridical perception is a matter of the absolute identity of type between the perceptual state of the believer and the actual properties of the object perceived. On one account[10] —this is Aristotle's position—when I see red I do so because some part of me (my eye-jelly) actually *becomes* red; and Epicurean perceptual theory commits them to something similar too. Such naïve realisms are easy targets, however, and no sophisticated causal theory of perception will be upset by such charges (although it is unclear just how sophisticated any ancient causal theory of perception actually was).

However that may be, it is likely that **290** principally directed against a rather different target. Atomists are committed to giving an account of every observable event as being the macroscopic phenomenal consequence of sub-perceptual atomic events, the rebounding and the occasional intertwining of the atoms themselves. The Atomists' explanatory ontology is deliberately parsimonious, consisting of the basic properties of weight, solidity, resistance, and the structural attributes conferred by the atoms' individual morphologies (VI.1*b*). The Fourth Mode fastens upon the basic atomist assumption that the fundamental microscopic properties will be no different in type from their observable, large-scale equivalents. The Atomists will of course allow that some macroscopic properties (colour, for example), are emergent, and have no direct micro analogues; but even so, their preferred fundamental trio of resistance, solidity, and weight are perfectly ordinary, everyday properties.

But what grounds have the atomists for thinking that they (unlike the Lockean secondary qualities) exist in their familiar form in the non- evident microscopic world? There seems to be no reason a priori (and even if there is, the Atomists do not offer it) why the real, atomic properties should not be utterly different in type from anything we come across in perception (as is indeed the case with the properties postulated by contemporary high-energy physics). The Atomists' assumption is, in fact, an analogue to naïve realism of perception, which one might label ‘explanatory naïve realism’; and Aenesidemus is perfectly justified in criticizing it as arbitrary and unjustified.

The Atomists might respond that we cannot conceive of properties utterly distinct from those which we experience, and hence cannot form theories on the basis of them. This may in a sense be right (witness the role

end p.274

of analogy in the development of quantum-physical concepts). But that fact, if it is one, cannot justify our ascription of these properties to the hidden world. Perhaps we simply cannot know what things are like down there at all. Or perhaps rather we can only comprehend them as causal hypotheses theoretically introduced to account for their observable effects, which do not, however, allow us to say what they are really like (compare the notions of Berkeley: *Principles* §§27, 142). Epicureans might still appeal to their theory's explanatory success in subsuming all observable phenomena under explanatorily prior atomic hypotheses. But even if atomism is explanatorily adequate in this way (as is at least questionable), the Sceptic may call upon the Second Mode: this sort of 'explanation' offers no warrant for construing the theory that provides it realistically (or for that matter even instrumentally, except in a provisional fashion).

But in any case, ancient atomism runs into insuperable problems in accounting for the operations of what we (and some ancients) would conceptualize as forces. In a typically polemical passage, Galen savages the inadequacies of the Epicurean atomist account of magnetism (*On the Natural Faculties* 2. 44–52; cf. Lucretius 6. 906–1089; XI.5*b*): a theory that attempts to explain it in terms of streams of hooked, interlocking, rebounding atoms somehow dragging things together across a distance is doomed to failure (equally, it cannot satisfactorily account for the transmission of magnetic power through a suitable medium).

At this point the Sixth Mode becomes relevant:

291 they frequently allow only such facts as are consistent with their hypotheses, while passing over those which conflict with it, even though they possess an equal persuasiveness. (Sextus Empiricus, *Outlines of Pyrrhonism* 1. 183 = 72m LS)

Dogmatists simply ignore anomalies, dealing only with those facts which their theories can account for. There is something in this. A single disconfirming instance falsifies a generalization; and if there are recalcitrant phenomena which simply cannot be explained by the Atomists, then atomism, at least as a universal hypothesis, is false. Yet Dogmatic practice may be defended here. Philosophers of science as divergent as Kuhn and Lakatos recognize that theories will inevitably run into evidential difficul- ties of this sort,[11] but hold that one should not precipitately abandon a

theory simply because it is in trouble. Provided that the anomalies are recognized as such, and can generate fruitful further theorizing, they need not be considered as malignant growths on the body of science. But whatever the truth of that, Aenesidemus' objection is a serious one.

The Sixth Mode is an empirical claim about the shortcomings of actual scientific practice, not a general, theoretical attack on the very possibility of science as such (as the First Mode appears to be). Similarly modest is the Fifth Mode,

292 according to which virtually all of them assign causes on the basis of their own hypotheses about the elements, and not on the basis of commonly agreed methods. (1. 183 = 72m LS)

The Epicureans postulate minimal atomic particles, and the existence of absolutely empty space. The Stoics and Peripatetics make the four elemental stuffs, or their constituent properties, fundamental, and believe that space and the matter it contains are continuous and dense. Platonists and Pythagoreans generate the world from numbers, lines, and planes, the limit and the unlimited. All of them produce aetiologies in line with their particular presuppositions, and there is no universally agreed, impartial method of deciding between them; hence they are simply incommensurable, and the partisans of the various competing theories are simply talking past one another. The Fifth Mode, then, supplements the first two, which establish that different explanations for the same phenomena are possible, by noting that the varying schools will simply pick arbitrarily upon whichever of those different explanations best fits their own theoretical prejudices. But matters may be even worse than that:

293 The Seventh [Mode] is that according to which they often give causes which are not only in conflict with the appearances but also with their own theories. (1. 184 = 72m LS)

It is not clear who (if anyone, in particular) Aenesidemus has in mind here, but once more the atomist account of magnetism could offer a tempting target, as also might the doctrine of the swerve, which is suggested by no macroscopic phenomenon and is in conflict with the principle that nothing occurs without a cause.

Finally,

294 The Eighth is that according to which, when things are frequently equally doubtful in regard to both those things which seem to be apparent and those under investigation, they construct their exposition concerning things equally doubtful from things equally doubtful. (ibid.)

The things which ‘seem to be apparent’ are descriptions of phenomena which are already implicitly slanted towards a particular explanation. Consider

end p.276

again the Stoics' favourite example, that sweating is an indicative sign of the existence of invisible pores (VI.4*a*, **246**). That supposedly secure inference from apparent to non-apparent rests crucially on the assumption that the surface moisture is indeed transpired from within the body. What masquerades as an evident fact actually requires interpretation: by describing the surface moisture as an emanation, one has already begged the question of its origins (against the supposition that, for example, it is precipitated out from the atmosphere). Interpreting the apparently evident antecedents in the sign-conditionals in this loaded fashion makes them ‘equally doubtful’; but only by treating them in this non-innocent manner can the Dogmatic theorists infer to the ‘equally doubtful’ consequence of the skin's microscopic perforations.

Dogmatists may still point to corroborating evidence: sweating increases after the ingestion of liquids, which suggests that it is caused by internal moisture. But these further considerations are themselves open to question, and none can make the inference logically watertight, contrary to the Stoics' explicit contentions (cf. arguments [A] and [B] of VI.4*a*). We can perhaps construct a web of explanatory beliefs which gains some plausibility from its overall coherence and explanatory power; but it will always fall short of deductive certainty. This much no doubt seems obvious to us, schooled as we are in the Humean limitations of empiricism. But it was not so obvious until Aenesidemus and his colleagues made it so.

The Aenesidemean anti-causal modes differ, avowedly so, in scope and power. The first, second, and perhaps also the fourth, make absolutely general claims about the theoretical ungraspability of the sort of 'truth' aetiological science professes to deal in: if they are sound, no realistic science of the non-evident foundations of things is possible. By contrast, the other modes point to contingent failings in the actual practice of scientists, failings which could, one supposes, in principle be eliminated. But taken together and in concert, as they were explicitly intended to be (*Outlines of Pyrrhonism* 1. 185), they form a powerful and coherent set of fundamental objections against characteristic Dogmatic practice.

3. The General Attack on Causes

(a) The Inconceivability of Causes

The Eight Modes of Aenesidemus are directed against scientific causal theorizing. But the Sceptics' general assault upon the concept of causing ranges much more broadly than that. Sextus Empiricus (fl. *c*.ad 200), our primary source for Greek scepticism, offers two sceptical treatments of causation. *Outlines of Pyrrhonism* 1. 13–30 gives a brief, schematic account
end p.277

of the views of varying Dogmatists, with sceptical replies to them. *Against the Professors* 9. 195–267 provides a more extensive, not to say rambling, compendium of arguments against the coherence of the notions of cause, effect, agent, and patient (see also IX.1*b*).[12] Sextus' shorter account opens with some general remarks:

295 as far as what is said by the Dogmatists is concerned, no one could even form a conception of cause, since as well as offering dissonant and incompatible conceptions of cause, they have made its instantiation undiscoverable by their dispute about it. For some say that cause is corporeal, others that it is incorporeal. But it would seem that in its general sense a cause is that because of which the result comes about. (Sextus Empiricus, *Outlines of Pyrrhonism* 3. 13–14)

The last sentence suggests that at the very least the Dogmatists concur over a very general characterization of what causes are; but their accord goes no further. They cannot agree, for instance, whether objects or their properties ('the sun', 'the sun's heat') are properly the causes of things; moreover, they quarrel over whether effects should be picked out by noun- phrases ('the melting of the wax') or predicates ('that the wax is melted'): 3. 14 (see VII.1*g*). And evidently the 'because of which' prepositional formula of the last sentence may be unpacked in a variety of ways: Aristotelians will take it to include final causes, while for the Stoics it will refer only to actual, active agents. Sextus elaborates upon the supposed inconceivability of causes:

296it is impossible to form a conception of the cause before apprehending its effect as its effect, since we only know that it is a cause of that effect when we apprehend it as an effect. But neither can we apprehend the effect of the cause as its effect if we have not apprehended the cause of the effect as its cause, since we seem to know that it is its effect only when we apprehend its cause as its cause. If, then, in order to conceive the cause we must first know the effect, while in order to know the effect we must first know the cause, the Reciprocal Mode [see §4 below] will show both to be inconceivable. (3. 20–2)

Sextus' procedure here is a species of a perfectly general type of sceptical argument, deployed to particular effect against Dogmatic accounts of sign-inference:

297it [i.e. the sign] will not be capable of revealing the consequent if the signified is relative to the sign and is for this reason apprehended along with it. For relatives are apprehended along with each other; and just as right cannot be apprehended before left as right of left, or the other way round (and similarly for all other relative terms), so it will not be possible for the sign to be grasped before the thing signified *qua* thing signified. And if the sign is not apprehended

before the thing signified nor can it be revelatory of that which is apprehended along with and at the same time as it. (2. 117–18)

Sextus' argument rests on a general assertion about relative terms (which will turn out to be equally applicable to the elements in the causal relation), to the effect that

(R) if *A* and *B* are correlative, then *A* cannot be apprehended before *B*.

But given that signs and what they signify are correlative, it follows from R that signs cannot be apprehended prior to what they signify. But in that case

298the sign is inconceivable. For they say that it is both relative and revelatory of the thing signified, in relation to which they say it is. So if it is relative to the thing signified, it should at any rate be apprehended along with the thing signified. . . . But if it is revelatory of the thing signified, it should at all events be apprehended before it, so that, being known before, it may lead us to a conception of the thing which is known from it. (2. 119–20)

If *A* is revelatory of *B*, then *A* must precede *B*; so if signs are revelatory of the things they signify, they must precede them, contrary to what has just been established. The whole notion of a sign is thus incoherent.

To rebut this argument, Dogmatists must either reject the notion that in order to be revelatory of its object a sign must precede it, or alternatively attempt to modify R. Sextus cites in support of R predicates like ‘to the left of’: if something is on the left, it must be to the left of something, which is, conversely, to its right; and the same goes for parents and children, the other stock ancient example of relatives:[13] (i) *A* is the parent of *B* if and only if (ii) *B* is the child of *A*. But (i) and (ii) do not express different facts, any more than ‘*A* signifies *B*’ means something different from ‘*B* is signified by *A*’; they are simply two different ways of expressing the same, relational fact (ancient accounts of relations were crucially vitiated by the tendency to treat relations, like everything else, as one-place predicates), a fact which has nothing to do with the relative dating of *A* and *B* them- selves. Sextus' account confuses the apprehension of the *items* which are the sign and what it signifies with the apprehension of the relation between them. As soon as I see smoke I know thereby that there must be fire; but it is not necessary that I see the fire at the same time; indeed I may never see it.

Applying this result to **296**, even if we cannot apprehend the cause as cause without apprehending the effect as effect, it does not follow that we

end p.279

have to apprehend each item both at the same time as and prior to the other (as Sextus alleges, and which is evidently impossible). We may distinguish two theses concerning the causal apprehension:

(CT1) Cause *C* is not apprehended prior to its effect;

and

(CT2) *C* is not apprehended *qua* cause prior to its effect.

The subject of CT1 is an item (e.g. an event); by contrast, the subject of CT2 is an item under a particular description, or, perhaps better, the relation between items in virtue of which that description can be seen to apply. Sextus' argument can now be seen to trade upon a confusion. Suppose causes (or at least some of them) precede their effects. In that case causes ought to be apprehended before their effects. But causes cannot be known to be causes before effects are known to be effects. But even if true, CT2 has no tendency to entail CT1.

Moreover, CT2 is in any case too strong. In general, in causal inferences, we are interested in relations between event-types: we come to form causal generalizations linking *A*-type events with *B*-type events (quite how does not concern us here). Thus we come to see (somehow or other) that a general relation (of dependence, or merely subsequence) holds between *A*-type and *B*-type events. When we see this, we apprehend the causal relation between them: we do not grasp first one and then the other. Thus

(CT2*) if *C* is the cause of *E*, then *C*'s being the cause of *E* is not apprehended prior to the apprehension of *E*'s being the effect of *C*

is trivially true, since the two clauses of the apodosis are equivalent descriptions of the same relation. But for all that, in individual cases we may yet first apprehend the particular *item C*: and hence (on the basis of CT2*) infer *E*'s necessary (or at least likely) subsequence. Sextus offers no adequate reasons for thinking that we cannot conceive of cause and effect at the same time.

(B) Causal Relations

Dogmatists dispute the proper linguistic categorization of causes (VII.1*g*, **271–2**; cf. **295**). More significantly, they disagree about what types of cause there are, and what role they play. Such disputes were not merely scholastic: they lay at the heart of a profound disagreement between Empiricists and various types of Dogmatic doctor as to the scope and function of explanation (see further IX.2). Moreover, we have already examined some

end p.280

of the various competing schemata for the classification of causes (VII.1*c*–*e*), much of our information for which derives from Sextus.

As was clear from discussion of the Stoic account (VII.1*d–g*), there are pressing reasons why any adequate causal theory must allow both for causes which precede their effects (*aitia prokatarktika*) and those which are concurrent with them (*aitia sunektika*, and their various concomitants): Sextus argues (*Outlines of Pyrrhonism* 3. 15–16) that you cannot have it both ways. Furthermore, some (cf. IX.1*c*) have rejected antecedent causes altogether on the grounds that

299 since the cause is a relative existent, and relative to the effect, it cannot precede it as cause, (3. 16)

for reasons analysed in the previous section. Still

300 either the cause produces the effect as already being and subsisting, or as not being a cause. Certainly not as not being; but if as being the cause must previously exist and come to be, and then brings about the effect which is said to be effected by it as already being a cause. But since cause is relative, and relative to the effect, it is clear that it cannot pre-exist as a cause. (3. 25)

Moreover,

301 the cause must either (i) co-subsist with its effect, or (ii) precede it, or (iii) come after it. But to say that a cause comes into existence after its own effect is laughable. But neither can it precede its effect, since it is said to be conceived along with it, and relative things, so they say, in so far as they are relative, coexist and are co-conceived along with one another. Nor can it coexist with its effect, since if it is effective of it, and what comes to be must do so as a result of something already existing, the cause must come to be earlier and then as such produce its effect. (3. 26–7)

The disjunction of (i–iii) is exhaustive; none of the disjuncts is coherent; consequently, neither is the concept of a cause (cf. IX.1*b*).

Sextus robustly dismisses (iii) as absurd. Some recent philosophers (not- ably Dummett 1978*a*, *b*) have been less hasty: but backward causation is at the very least conceptually problematic,[14] and it was not in any case taken seriously by any ancient philosopher. Moreover, Sextus' arguments against (ii) will, if successful, equally demolish the possibility of (iii), since they rely on principle R: cause and effect are correlative, and correlatives must coexist. But in the case of causes, there is a twist in the tail of the argument.

Consider again the dispute over the proper characterization of causes and effects (*Outlines of Pyrrhonism* 3. 14: **295** above; VII.1*g*, **268**). Even if we may, without incoherence, think of an *agent* as pre-existing any activity,

what about its causal properties? If it is *a*'s *F*-ness that makes *b F* (as PCS has it: I.3*c*), *a*'s *F*-ness should be causally effective for as long as *a* is *F*. But in that case, why is *b*'s *F*-ness not precisely coincident with *a*'s? For even where PCS may plausibly be thought to hold, the transfer of the property in question need not be (and generally is not) instantaneous: it takes time for the fridge to chill my champagne (of course further legitimate questions may be raised by the impatient as to *why* it takes time to do so).

Sextus' argument involving the relational status of cause and effect is more complex than the misguided argument from relativity canvassed in the last section might suggest. His case does not crucially rest upon the ontological co-dependence of relata generated by a naïve theory of relations. For causation is a physical relation, one thing producing another; and if *x* really is producing *y*, *y* should be there already. How then can *x* precede *y* as a cause? And when does *x* cause *y*? Barnes (1983*a*: 180–7) argues that the latter question is misplaced, since causing should be analysed as a two-place relation between events, each of which has the form Fx_t. Causal sentences thus take the form

(C) $C(Fx_{t1}, Gy_{t2})$.

Each related event is datable; but the relation *C*, not being an event, is not.

This is ingenious: but it leaves open some genuine questions. We can ask, without obvious solecism, when a cause is exercising its power; and if that turns out to be incoherent, ordinary language will need a fair amount of revision. Equally, we talk of processes progressing towards some conclusion, and here too temporal considerations seem appropriate. I may reasonably say that the sun is making me feverish after being exposed to it for a while, although as yet I have no fever, nor even any direct symptoms of fever. The important thing is to distinguish states from the processes which produce them: *x* is causing *y*'s *F*-ness as soon as it inaugurates the process which will ultimately culminate in *y*'s being *F* (it is worth noting that this is true even if *y* never, as a matter of fact, becomes fully *F*: interrupted processes must be characterized counterfactually in terms of the state they would have achieved had they not been interrupted). And *x* is doing the causing throughout the process. As soon as the champagne is placed in the fridge, the fridge is chilling it, even if it is still unpalatably tepid; and that will be no less true if, as a result of mechanical failure or a power cut, the champagne never attains a drinkable temperature.

This coheres well with the Stoic analysis (VII.1*g*). Antecedent causes are the temporally remote causes of events or states, which are best described by predicates ('that the champagne is chilled'); containing causes are the concurrent causes of processes, which are properly picked out by noun- phrases ('the chilling of the champagne'). Thus the Dogmatists need be

end p.282

guilty of no confusion or incoherence in claiming that causes both do and do not precede their effects.

(C) Action and Affection

Sextus has not, however, shot his bolt quite yet. In *Against the Professors* 9. 237–251 he runs through a set of arguments purporting to demonstrate the incoherence of any distinction between agent and patient. This is the first of them:

302(i) Furthermore, if a cause exists, it is the cause of something either autonomously and using only its own power, or else it requires assistance to this end from the matter affected, so that the effect is conceived of according to their mutual combination. (ii) But if it is naturally able to produce something autonomously and using its own power, it should, since it is always in possession of itself and its own power, always produce the effect, and not be productive towards some things and not towards others. (iii) But if, as some of the Dogmatists say, cause is not one of the absolute and independent things, but is rather among the relatives, since it is conceived in relation to the patient and the patient in relation to it, something worse will come to light. For if each is conceived in relation to the other, and of these one is productive and the other passive, there will be one conception but they will happen to have two names, of producer and the patient. And for this reason the efficient power will no more reside in it than in what is said to be affected. . . . (iv) For example . . . if fire is the cause of burning, either it is productive of burning autonomously and using only its own power, or it requires assistance to this end from the burning material. (v) And if it produces the burning autonomously, relying only on its own nature, it should, since it has always had its nature, always have been burning. But it does not invariably burn, but burns some things and not others; therefore it does not burn autonomously using its own nature. (vi) But if it does so because of the suitability of the burning wood, how will we be able to say that it, rather than the suitability of the wood, is the cause of burning? For just as when it is absent there is no burning, equally in the absence of the suitability of the wood, no fire takes place. (Sextus, *Against the Professors* 9. 237–43 = 72n LS; cf. **397**)

This is not unambiguous; and we need to distinguish two further sceptical theses about causation:

(CT3) if *x* is genuinely productive of *F*-ness it should always produce *F*-ness in what it is in contact with;

and

(CT4) if *x* is genuinely productive of *F*-ness it should be constantly *F*-producing.

CT3, which is clearly the weaker thesis, is suggested by the second sentence of **302**(v); elsewhere, on the other hand, Sextus' language seems to favour CT4. Indeed, one might argue on the basis of the arguments advanced that CT3 collapses into CT4, since the notion of cause entertained here is strongly non-relational; but contact is a relational property. In any case, even CT3 expresses an extremely powerful condition, and one which few if any candidates for causal status could hope to meet, at least if causes are construed as persistent agents, as the passage requires (Alcmaeon's soul is perhaps an exception: I.3*c*). But while Stoic containing causes are both necessary and sufficient for their effects (C4: VII.1*c*; Plato's proper causes are as well: P1–4: III.1*b*; cf. P7–8), they are not themselves agents or individuals, but rather processes or properties.

This argument is owed ultimately to the Alexandrian physician Erasistratus (IX.1*c*), who argued against antecedent causes on the grounds that they are insufficient for their supposed effects, a position rebutted at length by Galen (XI.2*b*): exposure to excessive heat is a cause of sunstroke, but not everybody so exposed falls sick, since it is not the only causally relevant factor in the pathogenesis; it is no part of the concept of causation that every item picked out as causally relevant should be constantly conjoined with its supposed effects.

But accepting that causing is co-operative, as Galen explicitly does, is precisely what, on Sextus' account, leads to something 'even worse' (**302**(iii)), namely the inability to distinguish in causal terms between the fire on the one hand and the wood's flammability on the other, since one is not actualized without the other. We will have to say that wood causes the fire to burn it just as much as the fire itself causes the wood to burn. But it is difficult to see what is so harmful about that: the causal *powers* involved still seem perfectly distinguishable, as are the triggering agents and the standing conditions that render their agency possible. It is no doubt true that we do tend to concentrate on one particular factor as the cause of an event to the exclusion of the others (the spark caused the fire, and not the availability of oxygen and flammable material);[15] and that predilection may cause us to overlook causation's co-operative nature, and hence foster some perennial ordinary-language causal fallacies. But we are in far worse trouble if we refuse to treat individual, non-sufficient causal factors as though they were causes: for that way we will mistake the actual causal structure of the world.

Sextus' next argument is in effect an extension of the previous one:

303 if there is a cause, it either has one effective power or many. . . . It has not one power, since if it had one power it ought to affect all things alike and not differently. For example, the sun burns the Ethiopian regions, warms us, but

end p.284

only illumines the Hyperboreans; it dries mud, but melts wax; it whitens clothes but blackens our skin . . . Consequently, if it had one power it ought to produce the same thing in all things. But it does not produce the same thing in all things. Therefore it does not have one power. But it cannot have many, since then it should effect all of them in everything, burning and melting and fixing everything, for example. (ibid. 9. 246–8)

The opposition counters, reasonably, that the effects produced vary from material to material (cf. IX.2*a*). Sextus replies that this is to deprive the agent of genuine causal power. It does so, of course, only on the assumption that causal powers must be non-relational facts about things. Causal powers are supposed to be genuine properties of things; and Greek metaphysics standardly considers properties to be genuine just in case they were non-relational. Hence if Sextus can make good the claim that causal properties are intrinsically relational, perhaps he can thereby show, contrary to Dogmatic pretensions, that they do not really belong to the object at all.

The solution here (as elsewhere in combating scepticism) is to develop a theory of dispositions; or equivalently to distinguish, as the Stoics did, between the category of relative things, and what they called 'relatives relatively disposed':

304 they [i.e. the Stoics] call (*a*) 'relative' all things which are conditioned according to an intrinsic character, but which are directed towards something else; and (*b*) 'relatively disposed' all those whose nature is to become and cease to be a property of something without any internal change or qualitative alteration. (Simplicius, *On the 'Categories'* 7. 166. 15–19 = SVF 2. 403 = 29c LS).

Examples of (*a*) are knowledge, sense-perception, and perceptual qualities like sweetness and bitterness. Instances of (*b*) are pure relations such as ‘to the left of’ (cf. §3*a* above). Only if causal powers are forced into category (*b*) will their status as genuine, intrinsic properties of objects be compromised: but the facts of relative effectiveness already adduced do not compel any such assignment.

4. The Modes of Agrippa

Nothing is known of this Agrippa, who is mentioned only once in the sources, in Diogenes Laertius' extremely brief sketch of the Modes (*Lives* 9. 88). Sextus, in his rather fuller account (*Outlines of Pyrrhonism* 1. 163–77), does not refer to him. The Modes are a supposedly complete compendium of refutational devices; and in so far as they bear on the Dogmatists' habit of reason-giving, they demand a brief treatment here. The First Mode alleges that dispute among Dogmatic theorists is universal and
end p.285

irresolvable. Every object of investigation is subject either actually or potentially to fundamental disagreement: but in default of any generally accepted or acceptable criterion for resolving them, no fruitful confrontation can take place between the opposing camps (1. 165; cf. **292**). The Second Mode (1. 166) points out that justification courts infinite regress: you assert *p* on the basis of *q*, *q* on the basis of *r*, *r* on the basis of *s*, . . . and so on; the process need never end. But if it does not end, then no real explanation has been given (as Aristotle for one agreed: V.1*a*). The Third Mode (1. 167) alleges that all assertions are relative in one sense or another, and hence that no genuine facts about things are recoverable. The Fourth Mode (1. 168) is that from Hypothesis: in order to evade the infinite regress threatened by the Second Mode, Dogmatists treat certain assumptions as basic; but they are merely assumptions, and hence can have no tendency to support whatever is deduced from them (cf. III.3*b*). Finally (1. 169) the Reciprocal Mode exposes circularities where the supposed explanans itself relies on the explanandum for confirmation, making neither well-founded.
The Five Modes clearly raise fundamental issues in regard to the status of the allegedly explanatory propositions in any science. As Sextus goes on to point out, taken together the Modes threaten to close off any attempt to justify a position, since either the process of justification will never end, in which case there is no justifier (Second Mode), or it will terminate in some arbitrary assumption (Fourth Mode), or it will move in a circle (Fifth Mode); and in none of these cases will there be any genuine explanation of anything.
The threat posed by the Five Modes may be countered in a variety of manners. We might, with some contemporary epistemologists, choose to defend a certain type of circular justification (provided the circle—or the ‘web of belief’ if you like—is sufficiently all-embracing). Alternatively, we might deny that any and every fundamental, non-derived proposition needs to be a mere hypothesis, a simple unsupported assumption (as Aristotle did: V.1*a*): perhaps it is an axiom which can just be seen to be true; or if that rationalist alternative proves unpalatable, maybe we can give an account of the basic principles of science as empirical hypotheses with predictive and explanatory power rather than

indubitably certain Aristotelian first principles, but not simply as unargued assumptions. Or one might think that explanation is indeed infinitely regressive without thinking that that in itself compromised the claims of each level legitimately to be explained by the next more fundamental stratum in the system (on the model, for example, of a recursive theory of truth). But at the very least Agrippa's Modes must make the Dogmatic scientist pause for thought, and for reassessment of the frequently all too comfortable assumptions upon which he has erected his explanatory edifice.
end p.286

5. Scepticism and Science

(a) Divination and Instrumentalism

We noted in the last chapter (VII.2*d*) the comfortable fit the Stoics found between their metaphysics and theology on the one hand, and a belief in at least the theoretical possibility of divination on the other. Indeed, a suitably strong notion of divine providence allied to the belief in universal *sumpatheia* might almost seem to entail that divination be possible:

305moreover, they hold that all forms of divination will be real if indeed providence exists. And they found the science on certain results, as Zeno says, and Chrysippus in the second book *On Divination* and Apollodorus and Posidonius in the second book of his *Physics* and in the fifth *On Divination*. Panaetius, on the other hand, declares it to be non-existent. (Diogenes, *Lives* 7. 149 = SVF 2. 1191; cf. VII.2*d*, **278**)

But support for divination was not confined to Stoic metaphysics; indeed, our sources attest to a bewildering variety of such practices throughout antiquity, perhaps the most important among them being divination by way of sacrificial entrails (hepatoscopy), by the flight and cries of birds (oionoscopy), and, at least from the Hellenistic period onward, astrology.[16] In addition, dream-interpretation was common from the earliest times. The Hippocratic *On Regimen* 4 deals with the importance of dreams for diagnosis and therapy; and Aristotle wrote a treatise, *On Divination in Sleep*, attempting to account for the apparent phenomenon. In the second century ad , Artemidorus published his *Dream-Interpretation*, a diviners' handbook, which none the less warns against the prevalence of charlatanry in the profession:

306prophets [are to be believed], but only such as do not deceive or mislead; for everything said by the Pythagoreans, the physiognomists, and the prophets who divine from dice, from forms or figures, palms, dishes, sieves, cheese, or necromancy must be regarded as false or misleading. These people's ‘arts’ are totally fraudulent, and they have no knowledge whatsoever of divination, but rather cheat, deceive, and swindle those whom they come in contact with. (Artemidorus, *Dream-Interpretation* 2. 69)

And such warnings of fraud are widespread (cf. Pliny, *Natural History* 30. 14, 19; Favorinus, in Gellius, *Attic Nights* 14. 1. 33). Moreover, such practices are notoriously prone to fraud; no one disputes the fact that the diviners’ predictions frequently turn out

false (on which see Cicero, *On Divination* 2. 115–16), even accounting for the legendary obscurity and ambiguity of oracles (ibid. 2. 117–18).
end p.287

The diviners have a number of answers to such charges, but they fall largely outside the scope of this inquiry. None the less, there is one such response which is of importance:

> **307** furthermore, someone who wished to buttress those who make prophecies might adduce as the reason for their not invariably lighting on the truth that not everything turns out in accordance with the nature of each individual and with fate, but some things happen contrary to it, while prophecies concern things which happen in accordance with fate. (Alexander, *On Fate* 7. 171. 7–11)

And although the Stoics would not have accepted the formulation that some things occur contrary to fate, they could well accept that divination deals only in general, natural tendencies of individuals which can be overcome by training or other factors.[17]
That serves to introduce an important distinction among the proponents of divination themselves. Some seers professed, in the manner of **307**, merely to be able to discern general tendencies of character from the individual's physiognomy (or horoscope, or whatever). Others, on the other hand, claimed to be able to descry the precise shape of the future. Elsewhere[18] (at least in relation to horoscopes, with which we shall be most concerned in what follows), I have called these positions ‘weak’ and ‘strong’ astrology respectively. The Stoic leader of the second century bc , Diogenes of Babylon,

> **308** conceded that they [i.e. astrologers] could foretell the general nature of each person, and what business they would be most suited to; but he denied there was any way of knowing the other things that they foretold; (Cicero, *On Divination* 2. 90)

that is, he allows weak, but rejects strong, astrology. More strongly, Chrysippus apparently accepted also the truth of certain astrological theorems (VII.2*d*). And most powerfully of all, the mathematical astrologers (or Chaldeans, as they were known, because of astrology's supposed connection with Babylon) held that an individual's entire fate could be read off from a suitably accurate computation of their birth-date.
These pretensions became the subject of sceptical attack. Carneades is known to have made a particular target of the claims of the diviners: at the beginning of his own refutation of divination in book 2 of *On Divination*,[19]
end p.288

Cicero recalls at length ‘the questions Carneades used to begin his discussions with’, which ask in rhetorical fashion what possible scope can there be for divination? If we wish to know about sickness, we talk to doctors; if astronomical knowledge is our goal, we seek out astronomers—and so on; equally, questions of physics, ethics, and politics are to be settled, if at all, by appeal to specialists in these subjects (2. 9–12). Divination seems to have no proper scope. A theoretical response on the part of the diviners is not difficult to concoct here: their knowledge of, say, medical prognosis is distinct from the doctor's precisely in that it is not (at least directly) derived from physical signs and symptoms. Moreover, such objections have no force against the claims of astrology (and other forms of divination) to supply a complete account of an individual's future fate, since that falls under the purview of none of the other sciences, even taken in combination.

But the bulk of the sceptical attack on divination takes a rather different tack. In *On Divination*, the issue between Quintus (the mouthpiece of the prophets) and Cicero himself (who adopts the sceptical position) turns fundamentally on the question of whether the diviners as a matter of empirical fact achieve a statistically significant level of accuracy in their predictions. The issue is complex, and fundamentally vitiated by the ancients' lack of any mathematically rigorous conception of probability, or the means of assigning it. But basically Quintus argues that, even though diviners from time to time get things wrong, their rate of actual success is too great simply to be accounted for by chance (1. 23), while Cicero counters that, once due allowance has been made for fraud, falsehood, and ambiguity both in the reporting of oracular success and in the content of the predictions themselves, there is no such 'success' to be accounted for at all:

309 how many of the events they predict actually occur? And if any of them does, what argument does anyone adduce to show that it wasn't by chance? . . . You know very well that things almost always turned out contrary to their prophecies. (2. 52–3)

Excessive reliance on flattering oracles was of no help to Pompey. People, furthermore, tend to recall the 'successful' cases, and forget the failures; thus

310 we may on occasion derive by guesswork from the innumerable delusions of drunken or crazy people things which seem to be prophetic: for who is there who cannot occasionally hit a target if he shoots at it all day? We sleep every night, and dream almost every night: what is surprising about our dreams sometimes turning out true? (2. 121; cf. 2. 66)

Cicero's treatment (particular in regard to the nature of probability) is sometimes naïve: but he sees clearly what the diviners need to be able to do

in order to bolster their claims to empirical predictive success; and equally clearly he sees that they are not in a position to do so. On the other hand, Quintus' arguments in favour of divination are not nugatory; and they introduce one very important notion into the debate over the nature of science. Cicero attacks the diviners because, among other things, they can give no causal account which will underpin their supposed success. Quintus, however, maintains, in sound empiricist fashion, that the mere fact that diviners lack a causal explanation for their success (if indeed they do) cannot of itself impugn their empirical successes (if of course there are any; 1. 12, 34, 70–2). Quintus' position is, then, instrumentalist: we should treat the predictions of the diviners empirically, and assess their methods only in so far as they provide effective predictive algorithms. It is enough if they merely 'save the appearances'.

The extent to which Greek science, in particular Greek astronomical science, was concerned with saving the appearances purely in an instrumentalist sense is a matter of some scholarly dispute,[20] and one to which we shall return elsewhere (XI.1*a*); but that Greek scientists from Eudoxus on were perfectly capable of offering what were, in effect, purely predictive algorithms with no pretension whatsoever to provide an accurate description of physical reality is not controversial; and Quintus places himself squarely in this tradition, one which also embraces the medical Empiricists (IX.2*a*–*b*).

The Stoics, of course, do think that they can essay at least a provisional explanation for divinatory success in terms of their notion of the universal causal interconnection of the cosmos, or *sumpatheia* (VII.1*b*). Cicero subjects this to an entertaining polemic (2. 29–34), and while conceding that certain phenomena (notably those involving the tidal effects of the moon: 2. 34; cf. X.2*a*; XI.1*b*) do exhibit such harmony, ridicules the notion that such harmonies could obtain in the way necessary to ground such practices as hepatoscopy:

311 one may cite countless such cases which indicate a natural connection between distant objects. Let us grant that; it does not rebut my contention here: how can any kind of cleft in a liver foretell financial gain? What natural conjunction, harmony, or association (which the Greeks call *sumpatheia*) can there be which links a cleft in a liver with my small profit, or my little income with Heaven, Earth, and the nature of things? (2. 34; cf. Favorinus, in Gellius, *Attic Nights* 14. 1. 2–4)

Moreover, Cicero objects to Quintus' instrumentalism as being scientifically inadequate, since it cannot (obviously) aspire to giving a proper causal account of things:

end p.290

312 I sought from you the explanation behind the whole of prophesying. But you took refuge in a remarkable hiding-place. Since you realized that you would be in difficulties when I asked you for the explanation of each type of divination, you said a great deal to the following effect: when you see the actual facts, you needn't seek explanation or cause—what happens, not why it happens, is what matters. As if I would grant that anything did happen, or that it was right for a philosopher not to inquire into why something happened! (*On Divination* 2. 46)

And while he is prepared to allow (at least provisionally) empiricist, instrumental accounts of such things as the efficacy of certain drugs, or of meteorological signs, since these are capable of long-term confirmation, such confirmations are unavailable in the unique cases of divinatory 'success' (2. 47).

(B) The Sceptical Attack on Astrology

We noted above the distinction between strong and weak forms of astrology; and the former is clearly more vulnerable to empirical refutation. Could all the Romans killed in the disastrous defeat at Cannae have had precisely the same horoscope, Cicero pertinently asks (*On Divination* 2. 97)? The Chaldeans, or practitioners of mathematical astrology, adopted the strong line; some of the Stoics (notably Diogenes of Babylon: 2. 90) held that horoscopes could merely indicate general tendencies of character and proper vocations; otherwise how could twins meet radically different fates (2. 97; cf. Sextus, *Against the Professors* 5. 90; Favorinus, in Gellius, *Attic Nights* 14. 1. 27–8; Pliny, *Natural History* 2. 29, 7. 160–5)? Cicero presses these charges (in arguments he attributes to Panaetius) at *On Divination* 2. 89–99, and they take two principal forms. First of all, he alleges, the astrologers by their own accounts must agree that it is a matter of empirical fact that people born under the same aspect meet different fates, while those born with different horoscopes live and die in the same way. But secondly he argues a priori that, while we might allow that a near neighbour like the moon might have detectable terrestrial effects, it is absurd to suppose that the planets, which are a great deal further away, could have any major impact on earthly matters at all (2. 91).

Of course, if they are to function merely as *signs* of future events, they need have no such effects; perhaps a providential divinity has set the heavenly bodies up in such a way as to provide indications of what will befall us (**305**; cf. 2. 101–2; Cicero, *Nature of the Gods* 2. 161–8; and see XII.2*d*). The notion of providence itself was the subject of a blistering sceptical attack by Carneades (*On Divination* 1. 4) and others (including the Epicureans: Epicurus, *To Herodotus* 76–7, 80–2; Lucretius, *On the Nature of Things* 2. 645–8, 1093–4; 5. 165–9); and Sceptics also attacked the
end p.291

view that it would as a matter of fact be beneficial for humans to know their fates (*On Divination* 2. 54–5, 105; Favorinus, in Gellius, *Attic Nights* 14. 1. 36). But many of the astrologers did indeed hold that the configurations of the heavenly bodies directly controlled terrestrial destinies:[21]

313 the seven stars [i.e. the sun, moon, and five inner planets] have the role of efficient causes of everything that occurs in life. (Sextus, *Against the Professors* 5. 5)

The fifth book of Sextus' *Against the Professors* is also known as *Against the Astrologers*: and it is a self-contained essay devoted to demolishing the Chaldeans' pretensions. Much of the material it contains is no doubt Carneadean in origin (and some of it is repeated in Gellius' report of the second century ad Academic sceptic Favorinus in *Attic Nights* 14. 1. 1–36). Sextus begins by summarizing, in confessedly amateur fashion, the doctrines of the Chaldeans (*Against the Professors* 5. 6–42). They divide the ecliptic into the twelve signs of the zodiac, within which they make further subdivisions; the sign rising at the moment of birth controls the individual's destiny, while the relative positions of the other heavenly bodies are also supposed to produce specific effects. All of this requires chronometric accuracy:

314 this is the method according to which they divide the circle of the zodiac into this many allotments. Similar to it appears to be the way according to which they say that they originally observed the horoscope of each birth. For at night, they say, the Chaldean sat on a high peak observing the stars while another sat with the woman in labour until she gave birth, and when she did so, he at once communicated it to the man on the peak with a gong. And when the latter heard it he noted the rising sign as the horoscope. By day, however, he attended to sundials and the motions of the sun. (5. 26–8)

Having outlined the system, Sextus goes on to the attack, seeking to undermine its very foundations (5. 49) by showing that the precise config- uration of the heavens at the instant of birth cannot be determined, which means that predictions of the type the Chaldeans claim to be able to make are impossible (5. 50–4). The actual time of birth cannot, he holds, be computed with any degree of accuracy (5. 65), while conception, which some astrologers take to be the crucial moment, is even less determinable (5. 55–64). And what are they to make of premature or late births (5. 66–7)? Even if these difficulties were overcome, the apparatus mentioned in **314** shows the impossibility of conveying immediately to the astrologer the instant of birth (5. 68–72).

Then again, the astrologers' own classifications are vague and imprecise: where does one sign shade into another? This and problems with accurate timing and preserving the absolute consistency required over a very long

end p.292

run of cases in order to build up suitable empirical correlations make the 'science' unobtainable (5. 73–80); and phenomena such as refraction further confound the data (5. 81–2).

Sextus invokes empirical considerations in order to show that the astrologers must either deal in extremely precise and restricted periods of time (otherwise they cannot explain how twins can have different destinies: 5. 88–93), which the earlier objections to their practice show to beyond their competence, or horoscopy is altogether a blunter instrument, in which case it is vulnerable to empirical refutation:

315 it is not reasonable that life is ordered according to the movements of the stars, or if it is so, it is at any rate inapprehensible by us. (5. 95)

Sextus concludes his treatise with a final remark on the difficulties that stand in the way of making astrology an empirical science:

316 In general, since they do not say that the stars tell them of the differences in men's lives, but rather that they observe them along with the positions of the stars, I say that if the prediction is to be firm, the same position of the stars must be observed not only once along with one individual's life, but a second time with a second and a third with a third, so that from the similarity of the outcomes of the results in all cases we might learn that when the stars adopt such-and-such a configuration, such-and-such an outcome will result. Just as in medicine we have seen that a wound to the heart is the cause of death having observed it along with not only the death of Dion, but of Theon and Socrates and many others, so too in astrology: if it is believable that this configuration of the stars is indicative of such-and-such a life, then it will at any rate have been observed not once in one case, but many times in many. So since the same configuration of stars is seen only after long intervals, as they admit, the recurrence of the Great Year[22] being once every 9,977 years, human observation will not traverse such epochs even in the case of one birth. (5. 103–5)

Purely instrumental 'soft' astrology can never be placed on any firm empirical foundation. Both Cicero and Sextus (not to mention Favorinus), then, bring powerful arguments to bear against the pretensions of some at least of the diviners. But the debate does not end there: and in later chapters we will see the considerable intellects of Ptolemy (XI.1*b*) and Plotinus (XII.2*d*) much exercised on opposite sides of the issue.

6. Conclusions

Greek scepticism developed in reaction against the excessively sanguine epistemologies to which the Greeks were prone (§1). Aenesidemus'

end p.293

Modes against the aetiologists (§2) comprise a set of general arguments against the Dogmatic practice of reason-giving and explanation. The arguments are of varying scope

and power; but taken as a set they at least succeed in showing that there are fundamental philosophical difficulties attending the attempt to investigate the hidden structures of things, as well as methodological shortcomings in the actual practices of some scientists. More generally still, Sextus seeks to show that the very concepts of cause and explanation are incoherent: causes and effects are relational (and hence inseparable), and yet one is supposed to precede (and be metaphysically prior to) the other (§3*a*–*b*). Furthermore, causes are supposed to account for their effects, yet no candidates for causal status appear to be sufficient on their own to produce anything (§3*c*). The general methodological Modes of Agrippa also pertain to causal explanation: any attempt to explain must be either circular, infinitely regressive, or arbitrary (§4). The sceptical arguments against divination in general (§5*a*) and astrology in particular (§5*b*) are important both for their successful undermining of pseudo-science, and, inadvertently, for their contribution to the understanding of what standards a genuine predictive and explanatory science must meet.
end p.294

IX Explanation in the Medical Schools

R. J. Hankinson

1. The Limits of Explanation

(a) Diocles of Carystus

It was already apparent with the Hippocratics that medicine and philosophy enjoyed a close and mutually supportive relationship. In particular, medical theory made significant contributions towards the growth and nourishment of the various empiricist strands in Greek scientific epistemology. It is to their later development that we now turn.
The following long fragment is from the Sicilian doctor Diocles of Carystus (fl. *c.*325 bc):[1]

317[A](i) those who suppose that things [sc. foodstuffs?] which have similar juices or smells or degrees of heat or anything else of this sort have identical powers are mistaken, since one can show that many dissimilar things result from things which are similar in this way. (ii) Therefore none of these things should be supposed to be laxatives or diuretics or to possess any other power on the grounds that it is hot or cold or salt, given that not everything sweet, or acrid, or salt, or anything else of this sort, possesses the same powers; (iii) rather its whole nature should be thought to be the cause of whatever is accustomed to result from each of them; this way one will be least likely to miss the truth. [B](iv) Those who think that one should state a cause in every case . . .[2] seem to be unaware first that it is not always necessary to do so from a practical point of view, and (v) second that many things which exist are somehow in their nature akin to principles, so that they cannot be given a causal account. Furthermore, (vi) they sometimes err in assuming what is unknown, disputed, and implausible, thinking that they have adequately given the cause. (vii) You should disregard people who aetiologize in this manner, and who think that one should state a cause for everything, and rely instead upon things which have been excogitated over a long period on the basis of experience (*empeiria*); (viii) and you should seek a cause for those things which admit of it when it will make your account more understandable and

end p.295

more believable. (Diocles, in Galen, *On the Powers of Foodstuffs* 6. 455–6,[3] = Fr. 112 Wellmann)

317[B]appears Aristotelian in tone; and Diocles has sometimes been painted as a straightforward follower of Aristotle, his rough contemporary.[4] But although he says that some things cannot be assigned a cause (**317**(v)), Diocles does not commit himself to the view that these really *are* first principles: rather they resemble them in being inappropriate subjects for causal explanation. His claim is not, then, Aristotle's that science must be founded upon secure, indemonstrable first principles (V.1*a*), but rather that nothing practical is to be gained from trying to give aetiologies for everything (**317**(iv)). Explanation must indeed stop somewhere, but for pragmatic rather than metaphysical reasons: practical medical epistemology demands that some things be regarded as basic even though they may in fact not be (compare Herophilus: §1*b*, **320**).

317(vi)prefigures the Fourth Agrippan Mode of Hypothesis (VIII.4); but its point is implicit in Aristotle, and perhaps also in Plato (*Republic* 6. 510c–11d: III.3*b*). Diocles does not object to causal hypotheses as such, but only to those which are ill-founded and over-confidently asserted. This recalls *Ancient Medicine* (II.1*d*): the scientist should avoid rational constructions empty of empirical content, dealing only with propositions derived from experience; but crucially for Diocles, no less than for the author of *Ancient Medicine*, that need not rule out causal explanation altogether (although (viii) perhaps hints, more radically, that aetiologizing may be of rhetorical use only; compare the Hippocratic attitude to prognosis: II.1*b*).[5]

Yet as part [A] of **320** amply demonstrates, Diocles is not simply a sceptic in regard to explanation. He objects to the explanatory practices of some of his opponents; but he

does so on the basis of a positive view about the structure of explanation. The opponents (here perhaps including the author of *Ancient Medicine*, as well as *On Regimen*)[6] are prone to infer, apparently, that if one particular food is (say) both sweet and laxative, then anything else sweet will be laxative too: but this is empirically refutable (**320**(i)); and hence the phenomenal properties cannot be the causes of the powers either.
end p.296

Thus Diocles effectively endorses Plato's principle P2 (the same cause cannot produce opposite effects: III.1*b*), and not merely dialectically (as a Sceptic might) in order to undermine the coherence of the concept of causation itself. He happily recognizes, without sceptical qualms, that substances have powers or faculties: he was the author of an influential text on *Affections, Causes, Treatments*. What he rejects is the oversimplified and schematic approach of some of his contemporaries: in forming a just assessment of the case one must look not to one factor alone, but must consider everything of causal relevance: it is thus 'the whole nature' of the foodstuff which is likely to be responsible for its particular set of powers. Diocles emerges as a causal theorist of some sophistication, whose pronouncements on the matter foreshadow Galen's much more detailed variations on the same theme (§1*c* below; XI.2*b*).

Diocles' fragment further suggests (although it does not develop) the rudiments of an empiricism more thoroughgoing than Aristotle's. For Aristotle the realist, the one true explanation is always in principle discoverable, and, with a little effort and luck, practically recoverable by the assiduous scientist. Diocles is not convinced that such explanantia, even if they exist, will yield to human investigation; and if that is right, there may be no way of determining, a priori and without practical experience, what the powers of anything in general (or foodstuff in particular) are going to be.

(b) Herophilus

The third century bc witnessed the remarkable rise of Alexandria as the scientific and cultural centre of the Greek world.[7] Only a generation after its foundation by Alexander the Great it became the capital of the Greek kingdom of Egypt, while the patronage of a succession of Ptolemys ensured its rapid growth as the cultural capital of the Mediterranean world, soon surpassing an Athens in political and social decline. The focus of this intellectual efflorescence was the great library, but the culture nurtured scientists as well as literary men, pre-eminent among whom was Herophilus of Chalcedon (*c.*325–*c.*255 bc).[8]

Herophilus' principal claim to scientific immortality rests upon his anatomical discoveries. He distinguished between, and correctly described, the several functions of the sensory and motor nerves (T 80–1 vS[9]); he
end p.297

performed detailed dissections of the eye (T87 vS) and the brain (T 75–9 vS); and he made considerable advances in reproductive (T 101–14 vS) and vascular (T 115–28 vS) anatomy. His physiology and pathology were less innovative (he apparently subscribed to some version of the humoral theory: T 130–2 vS),[10] although he made contributions to pulse theory which established sphygmology as one of medicine's central diagnostic tools

(T 144–88 vS). This striking new knowledge was won, if the tradition is correct, by ethically dubious means. Celsus (a medical encyclopedist of the first century ad) reports that Herophilus and Erasistratus undertook vivisections of criminals from Ptolemy's prisons (*On Medicine*, Proem 23–6 = T63a vS; cf. 63b–c, 66 vS). This story has often been questioned, but usually on the inadequate grounds that no great scientist could behave in so beastly a fashion. Celsus offers a utilitarian argument in favour of the practice: the deaths of a few, who will in any case die unpleasant judicial deaths, may in the long run save countless others (ibid. 26). At any rate, Herophilus, unusually for antiquity, almost certainly dissected human cadavers (T 64–6 vS), which underscores the empiricist cast to his epistemology.

His attitude towards causal theorizing is summarized in two fragments:

318 the cause, whether or not it exists, is by nature undiscoverable; but in my opinion I believe I am chilled, warmed, and filled with food and drink (Herophilus, in Galen, *On Antecedent Causes* 16. 198[11] = T 59a vS)

319 some, such as Herophilus, accept causes 'on the basis of a hypothesis'. (ibid. 13. 163 = T 58 vS)

The import of these fragments is disputed.[12] But whatever 'causal hypotheticalism' amounts to, Herophilus clearly insists that the real metaphysical structure of the world is something irrecoverable (at least with certainty) by human investigation. Yet even if those structures must remain for us in perpetual obscurity, we may yet construct a perfectly serviceable account of the world's workings on the basis of the appearances, or *phainomena*:

end p.298

320 let the *phainomena* be said [to be] first (*prōta*), even if they are not first.[13] (Herophilus, in Anonymus Londinensis, *Iatrica Menonia* 21. 22[14] = T 50a vS; cf. Galen, *On the Therapeutic Method* 10. 107 = T 50b vS)

Von Staden (1989: 125) translates 'let the appearances be described first even if they are not primary': but the point seems rather to be that we must treat the appearances as basic (in the construction of explanation), even though there may well be, as a matter of fact, deep metaphysical truths which underlie them, since the latter are irremediably unavailable to us; and any provisional inferences we may attempt to make to them are only as secure as the *phainomena* upon which they rest.[15]

This sheds light upon **318** and **319**: what we have to go on are appearances, and any causal hypotheses invoked to explain them can never be more than hypothetical. Other fragments (T 52–4 vS) exalt the importance of experience and observation; and Herophilus may well have seen himself as following in the tradition of Aristotelian empirical science (cf. *Parts of Animals* 1. 5. 644^{b}32–645^{a}36: one must get one's hands dirty in the interests of science).

Even so, **318** seems more uncompromisingly sceptical in tone: it is intrinsically impossible to demonstrate any cause for a particular effect; and, if Galen's report is to be trusted, it becomes merely a matter of opinion. First appearances notwithstanding, **318** does not assert that Herophilus doubted even the reality of the subjective states of being warmed, etc.: rather, as the last example indicates, what is doubtful is the role played in their generation by other facts and events. I eat, and feel full: I believe that one causes the other, but strictly speaking I am not, and cannot be, directly aware of that fact. Herophilus, then, is not committed to any generalized doubt regarding the existence of causes as such—he can perfectly well allow that there is causal activity. Rather he questions our ability to be certain, for any candidate instance, about what precisely is responsible for it. Herophilus thus anticipates Aenesidemus' aetiological Modes (VIII.2). No amount of empirical evidence can entail the truth of any particular explanation; explanations are at best provisional.

This squares well with Herophilus' evident ease elsewhere with the giving of causal explanations. He was quite happy to postulate four 'powers' (*dunameis*) as being regulative of all living creatures (T 131 vS: the actual details of his view are obscure); and he is praised by Galen

end p.299

(*On the Formation of the Foetus* 4. 678–9 = T 57 vS) for realizing that descriptive anatomy alone cannot provide the inquirer with an understanding of how the parts actually function (cf. IX.5*a–b*). Pliny says that he 'inquired into the causes of disease' (*Natural History* 26. 8. 14 = T 209 vS). He also discussed a 'vital power', presumably a life-force of some kind (T 164 vS); and he allowed that some types of motion are 'discoverable by reason' (T 142a–c vS). And while he apparently[16] held that some fevers arose 'with no preceding cause' (T 217a vS), or at least with no apparent one (T 217b vS), he was none the less happy to ascribe certain cases of sudden death 'with no apparent cause' to paralysis of the heart (T 212 vS).

These are not the remarks of an out-and-out causal sceptic (compare the attitude of the Hippocratic *Epidemics* to visible occasions of disease: II.1*b*, **73–4**; cf. **63**; and cf. Diocles: § *b*, **320** above). On the other hand, talk of powers and faculties is deliberately vague: a 'power' is simply a placeholder for whatever it is that is, as a matter of fact, responsible for the effect in question. That sounds dangerously like Molière explanation,[17] and in careless hands can easily become so. But, as Galen realized (**422**: *On the Natural Faculties* 2. 1–10; XI.5*b*), provided one does not mistakenly imagine that explanation simply consists in the postulation of such powers, they can fulfil a perfectly legitimate methodological function: they encourage and give shape to the search for what precisely these functionally defined powers consist in. Herophilus, then, is not a sceptic with regard to causal explanation; he is simply concerned that the human desire to explain does not outrun its own limited resources.

None the less, this happy construction of Herophilus the sophisticated methodologist must confront one apparently recalcitrant text. In the passage immediately preceding **318**, Galen accuses Herophilus of pusillanimity:

> **321** one should be amazed at him [sc. Herophilus] and rebuke him for timidity—for, having called every cause into question by means of many powerful arguments, he himself is discovered making use of them, saying that they appear to be thus to many people. (Galen, *On Antecedent Causes* 16. 197 = T 59a vS)

Moreover Galen (16. 199–201) ascribes the argument of **301** (VIII.3*b*) against the coherence of the ordinary concept of causing to Herophilus, along with two other arguments of a similar structure also to be found in the central passage of *Against the*

Professors 9 (210–36) presenting the sceptical case against causes. The first goes as follows:
end p.300

(1) if there are causes, then either (*a*) bodies cause bodies, or (*b*) incorporeals cause incorporeals, or (*c*) bodies cause incorporeals, or (*d*) incorporeals cause bodies;

but

(2) neither (*a*), nor (*b*), nor (*c*), nor (*d*);

so

(3) there are no causes.

(1) rests on some such premiss as

(4) everything is either a body or incorporeal;

given the more or less evident truth of (4), (1) becomes an analytic truth (at least, if the logical form of causal propositions it presupposes is acceptable).

The arguments for each of the negative disjuncts of (2) are not given by Galen. They are, however, rehearsed at by Sextus at *Against the Professors* 9. 214–16, and there seems no reason to doubt that Herophilus would have offered something along the same lines. That said, they are not all that impressive. Body cannot be the cause of body (it is alleged), since cause and effect cannot have the same nature; (*b*) is disposed of by parity of reasoning; while (*c*) and (*d*) are rejected on the grounds that contact is conceptually basic to causing, and that contact can only occur between bodies (essentially held by Stoics, Epicureans, and indeed Aristotle: cf.
VII.1*a*). The weakest link is clearly the argument levelled against (*a*). It fails to take account of the possibility of potentiality (perhaps all bodies *are* potential causes; but they need not always actualize that particular potentiality); and it rests implicitly on some such principle as

(PP) for all *x* and all *y*, if *x* and *y* are members of some class *C*, then for any *F* such that *F* is a property of *x*, *F* is also a property of *y*.

PP is massively implausible for an unrestricted domain of properties—but even if it is restricted to essential properties of that class only (in which case it becomes less implausible), the argument founders on the considerations of potentiality. A second, more sophisticated, set of arguments (attributed to Aenesidemus) is rehearsed by Sextus at *Against the Professors* 9. 218–26, relying essentially on a version of the view that effects and causes must resemble one another in structure (PCS: I.3*c*); but even these ultimately do not threaten any moderately sophisticated account of the causal relation.

But, irrespective of their strength, what could Herophilus have been doing by invoking these arguments in the first place? It is one thing to say
end p.301

that causal ascriptions are inherently fallible, and should be accepted only 'on the basis of a hypothesis', quite another to attack the very coherence of our ordinary causal notions. Indeed, the second manœuvre seems to undermine the first. Perhaps Herophilus merely wished to draw attention to the internal difficulties experienced by any Dogmatic account of the causal relation—this would show not necessarily that no such account was feasible, but simply that none had yet been produced which was satisfactory—and that in turn might be construed mildly as showing only that we need abandon our pretensions to causal infallibility, and to an understanding of the arcane mechanisms whereby causal power ramifies. That there are causes is, in one sense at least, one of the *phainomena* (it seems that there are to most people); but there the general agreement ends. But that is not an entirely satisfactory explanation of the mismatch between the differing reports of Herophilus' position: there may well be something to Galen's charge of inconsistency.

(c) Erasistratus

Herophilus' contemporary Erasistratus is also linked with the practice of human vivisection (§ *b* above). They are usually thought to have been Alexandrian colleagues; and while other texts connect Erasistratus with the Seleucid dynasty in Antioch, there seems no good reason to doubt that he also worked for a time in Alexandria.[18]

He too was known for his anatomical skill and discoveries (the distinction between sensory and motor nerves is also attributed to him: T 81 vS = Fr. 39 G[19]). His basic physiological concept was that of the *triplokia*, the triple plaiting of three basic types of vessel (nerve, artery, and vein), which he held to be the basic elements of all bodily tissue, and from which the organic parts were developed.[20]

Galen, our principal source for his views, is characteristically hostile to Erasistratus, ridiculing his abandonment of the 'natural powers' of attraction and repulsion as explanatory tools in favour of the materialist principle of *horror vacui* as the sole agent of internal movements of material in bodies.[21] Moreover, as the proponent of a universal teleology himself (XI.3*a*), Galen rejects as absurd Erasistratus' Aristotelian claim that some organs are functionless (Fr. 81 G). Even so, Erasistratus was not opposed to teleological explanations as such (he considered their discovery to be the proper business of the philosopher: Frs. 83, 114 G; cf. Frs. 77–8 G); but

end p.302

he considered teleological explanation to be of strictly limited applicability—not everything in the world could be explained as subserving some purpose. In this he resembles Theophrastus more than Aristotle (V.4*a*), and this may be one source of the tradition (reported in Diogenes Laertius, *Lives* 5. 57: Fr. 7 G) that Erasistratus was Theophrastus' pupil.[22]

But in many respects, as Galen is not slow to point out, Erasistratus rejected utterly the Peripatetic picture. He abandoned the humoral theory of human constitution (Fr. 92 G). In his pathology he dismissed the notion that alterations of temperature had anything to do with disease, holding rather that all fevers are subsequent upon inflammations caused by transfusion (*paremptōsis*: Fr. 109 G) of the blood from the veins to the arteries, which (according to Erasistratean physiological principles) should properly contain only *pneuma*. Inflammation consists in *pneuma* displaced and compressed by the pathologically infused blood.

But for Galen at least (whose treatise *On Antecedent Causes* is mainly directed against the Erasistratean position) Erasistratus' greatest mistake in pathology was his complete disregard of the importance of external heating and chilling upon the human body. Heat and cold are, for Galen, among the most important pathogenic factors, and almost always figure prominently in the lists of antecedent causes he records (XI.2*b*);[23] yet Erasistratus considers them responsible only for superficial alterations in animals' conditions, having no impact on their internal dispositions (Fr. 75 G; cf. *Ancient Medicine*: II.1*d*).

This difference of opinion is of some theoretical importance. Erasistratus holds that heat and cold *cannot* be the causes of illness, since they are not invariably followed by it, and do not persist at the time of the illness.[24] For Erasistratus, then, if something genuinely is a cause of something else it must be invariably conjoined with it, and it must actually be in the process of effecting it. In other words all causes will be, for him, *aitia sunektika* (VII. *c*–*d*):[25] nothing else will meet the criteria for being a cause. The bulk of *On Antecedent Causes* is concerned with refuting this view (XI.2*b*).

But it is one thing to say that no genuine cause can precede its effect, or fail on occasion to be associated with it; quite another to adopt the view that no item, unless constantly conjoined with some other, can have any causal relevance to it. Galen sometimes tries to pin the latter view on Erasistratus; but it is unclear that Erasistratus needs to accept this
end p.303

consequence. He allows that overeating and exhaustion (stock later examples of antecedent causes) may figure in the onset of disease, although in line with his terminological strictures he apparently refused to allow them the title of ‘causes’.

Crucial to the Erasistratean pathology of fever is *plēthōra*, congestion of the veins caused by an influx of undigested food into the vascular system (Fr. 161 G). When the body ingests too much food the digestive system cannot cope with the overload, and, if evacuation does not take place by other means, the nutriment enters the veins, compressing the blood and forcing it from the veins via the capillaries into the arteries, causing inflammation and fever. Yet *plēthōra* is not invariably followed by disease, since the condition can be treated by evacuation and the pathogenic process arrested before the disease becomes established (Fr. 162 G). But once the transfusion, *paremptōsis*, has taken place, disease inevitably follows.

As Galen presents it, Erasistratus holds that only the *paremptōsis* can properly be called the cause of the illness, since only when that has occurred is the pathological outcome unavoidable. Erasistratus in effect stipulates that causes must be at least sufficient for their effects (it is less clear whether they need be necessary for them as well); and, since it is allowed on all sides that such antecedent factors as heating and chilling are not invariably attended by disease, they cannot be accounted causes:

322(i) thus the Sophists find reasons for their arguments in their attempt to show that, even if on some occasion these things harm weak bodies, not even then can they properly be called causes. (ii) For if indeed it is the case that they act because of their own internal nature, and that this action derives from themselves, and not from the weakness of the bodies, then they must be seen to have an effect at all times. (iii) But this is not in fact the case: things are rather as we observe them in the theatre. For it can happen that four out of a thousand spectators develop a temperature, and of these only one becomes feverish rather than all of them. (Galen, *On Antecedent Causes* 1. 9–2. 11; cf. 8. 96–115)

Galen does not name his 'Sophists' here; but later he attributes both the theoretical stipulations of (i) and (ii) and their empirical underpinning in (iii) to Erasistratus (8. 104; 10. 126–30) and his followers. It is worth reproducing Galen's account of Erasistratus' description of the aetiology of fever:

323he says something along these lines: 'This much, namely that there would not be a proximate cause of disease had something not first brought about something else, and then that something else again, no one denies. Congestion in the veins must follow repletion from things that have been eaten and drunk; and *paremptōsis* (i.e. a falling to the arteries) occurs because of this (let us allow the present argument to proceed in accordance with his dogmas); and in the course of the *paremptōsis*, the heart changes the *pneuma* more powerfully

and rapidly, which invariably propels the displaced blood and forces it to the edges of the arteries, where it is compressed and forms an inflammation. So it is clear that *paremptōsis* occurs by way of *plēthōra*, and then the compression occurs.' But on this account, even if the repletion is not the proximate cause of the inflammation, he none the less accepts that it is first in the list, and earliest, of the causal sequence just mentioned. And he says that *plēthōra* (as he calls vaso-congestion) comes about because of repletion, and *paremptōsis* because of that, and then he agrees that the congestion and the inflammation come about because of the *paremptōsis*, but he does not want to concede that repletion is the cause of illness, because of a dispute about nomenclature. This shamelessness is even worse than the others. Repletion causes *plēthōra*, *plēthōra* causes *paremptōsis* of the blood to the arteries, and *paremptōsis* causes inflammation; yet even given all this, he won't agree that repletion is in any way the cause of inflammation. (14. 174–6; cf. Fr. 198 G)

Erasistratus, on Galen's account, denies the transitivity of the causal relation, while holding that only proximate causes are genuinely causal (cf. VII.2*d*). Erasistratus evidently held unorthodox views both on pathology and on proper causal classification, but it is unclear whether the primary thrust of his causal revisionism was substantial or merely linguistic. It is, as Galen points out, one thing to say that there is something wrong with *calling* these non-sufficient antecedent conditions causes (although, as Galen remarks elsewhere, such stipulations fly in the face of ordinary Greek usage); it is something else altogether to suggest that nothing which fails to meet the strict criterion of causal sufficiency is even causally relevant. Yet it seems that Erasistratus was tempted to say the latter:

324 Erasistratus himself held that diseases were not produced by such causes [sc. heat, cold, insomnia, excessive indulgence in food, drink, or sex] since other persons, and even the same person on different occasions, was not made feverish by them. (Celsus, *On Medicine*, Proem 54)

Celsus himself rejects the position in a manner later elaborated upon by Galen (XI.2*b*):

325 the actual evidence is against his view, since diseases rarely occur except following one of them; and in any case it does not follow that what does no harm to some patient on one occasion may not harm some other, or even the same one on another occasion. (ibid. 58)

Celsus urges the claims of antecedent conditions to be causes on the grounds that they are (virtually) necessary conditions of their outcomes; and in any case a failure of universal efficacy on the part of some alleged causal factor *C* does not in itself show that *C* is utterly devoid of causal power. The dispute here parallels a similar debate in the case of sign theory. Aenesidemus argued that there could be no such thing as an evident sign of something, since people disagreed as to the interpretation

end p.305

of any candidate (indeed, his example is of medical disagreement: **335** below); and such arguments were taken over to their own ends by the medical Methodists (§3*a*, **342–3** below).

Celsus thinks the error jejune, and indeed holds that Erasistratus contradicts himself:

326 if he had had a sufficient understanding of the nature of things (which such doctors rashly arrogate to themselves), he would have known that nothing at all comes to be as a result of one cause alone, but what seems to contribute most to the effect is taken to be the cause. Moreover, it is possible that what alone has no effect may prove most efficacious when joined with others. Furthermore, not even Erasistratus himself, who says that fever is produced by the transfusion of blood into the arteries, and that this occurs in an over-replete body, can explain why of two equally replete people the one falls ill while the other avoids all harm—and yet this happens every day. (ibid. 59–60)

Whatever his actual position (he seems to have been happy to call *plēthora* at least an *archē* of fever, although not its cause: Frs. 162, 223 G;[26] but it is not clear what he would call repletion), Erasistratus evidently exercised a considerable influence on subsequent causal theory, one strong enough for Galen to consider it worth rebuttal some 400 years later.

2. Empiricism and Rationalism

(a) The Roots of Empiricism

Empiricism, in a general sense, was part of Greek medicine from earliest times. But in the generation following that of the great Alexandrians, there arose a school of doctors who

actually labelled themselves Empiricists. The school's origins are obscure; the medical *Introduction* falsely attributed to Galen traces the school back to the fifth-century Sicilian Acro: but this is a typically fanciful piece of late antique genealogy.[27] Better candidates for founder are Philinus of Cos (fl. *c.*250 bc): but virtually nothing is known about him (he is mentioned by Galen, *Outline of Empiricism* 1 = Fr. 10b, 43 D); and Herophilus' pupil Serapion (fl. *c.*225 bc).
end p.306

Medical Empiricism grew out of the small ‘e’ empiricism we have observed in Diocles and Herophilus, in its emphasis on the importance of observation and caution in aetiologizing (§§1*a–b*). The Empiricists went a stage further: for them, causal theorizing had no role to play whatsoever, restricted or otherwise. They confined themselves to dealing in more or less highly confirmed concatenations of evident events, for which they developed a sophisticated epistemology, both positive and negative. Medical Empiricism had an active history of some 500 years; but it never became an ossified orthodoxy. Ample evidence survives of lively debates within the school concerning the proper structure of medical science, debates that ebbed and flowed with the philosophical currents of the times.

Indeed from this time onward the history of philosophy and that of medicine become even more closely intertwined. Our principal source for later Pyrrhonian scepticism, Sextus Empiricus, was, as his handle suggests, an Empiricist doctor; and he was by no means the only person to practise both philosophy and medicine. The list of Pyrrhonian philosophers appended by Diogenes to his treatment of scepticism (*Lives* 9. 115–16) includes several people known to have been doctors (Menodotus, Theodas, Sextus himself, Saturninus), and others who may well have been.[28]

(b) The Empiricist Account of Medical Science

The most striking feature of the Empiricists' position was their consistent refusal to let their theorizing take them beyond the realm of immediate experience and into the arcana of things by nature obscure, the ‘hidden things’ or *adēla* of later Greek epistemology. That is, in the terminology of the parallel debates among the philosophical schools, they refuse to countenance indicative sign-inference (VI.4). However, congruently with the Pyrrhonist position, they happily allow the type of sign Sextus calls ‘commemorative’, and to build up general ‘theorems’ on the basis of them.

The Empiricist observes that affections (*pathē*) arise in people, sometimes for no evident reason (such as nosebleeds), sometimes with some apparent antecedent cause (as when bleeding follows upon a fall or a wound). Moreover, it is equally a matter of observational experience that in some cases interventions produce (or are at least followed by) beneficial or harmful results. Sometimes, Empiricists hold, we are naturally driven by a species of psychological compulsion to test out some remedy; sometimes something beneficial happens by chance; and sometimes we simply try anything at

random, improvisationally, in the hope that it may happen to be efficacious. If an improvement in the condition seems to follow the
end p.307

intervention, however it was suggested, we may employ it again in similar circumstances:

327 an imitative experience is one where something which has proved to be beneficial . . . is tried out again for the same complaint. It is this kind of experience which has contributed most to the art: for when they have imitated, not two or three but very many times, what has turned out to be beneficial on earlier occasions, and when they then find out that, for the most part, it has the same effect in the case of the same diseases, then they call such a recollection a theorem. (Galen, *On Sects* 2, 3 H[29])

The Empiricist, then, builds up a collection of observations where particular types of event are seen to follow one another (which he calls an 'experience', or *empeiria*); if it becomes sufficiently large it will generate a general rule, or theorem.

These theorematic relations between observable events need be neither universal nor positive: the Empiricists employ a fivefold typology of connection and disjunction, according to whether things are seen to go together always, for the most part, half the time, rarely, or never (*Outline of Empiricism* 2, 6 = Fr. 10b, 45–6, 58 D). These categories stand in determinate logical relations: always *p* if and only if never not-*p*; for the most part *p* if and only if rarely not-*p*; half the time *p* if and only if half the time not-*p*. And all of them are of value in isolating appropriate therapies and rejecting others.

Personal observation, *autopsia*, lies at the bottom of the Empiricist doctor's *empeiriai*; but the good Empiricist will also take account of the reports of others, *historia*, in order to settle on a particular therapy in a particular case. Not that they simply accepted second-hand intelligence uncritically: they were aware of the intrinsic unreliability of hearsay, and contrived a sophisticated theory of evidential acceptability which weighed both the previous reliability of the particular source and its agreement with others.

Indeed the Empiricists self-consciously applied their own epistemological standards to the assessment of the reliability of the reports of other doctors (8 = Fr. 10b, 65–9 D); and they carefully avoided being lured into any Dogmatic claims regarding any particular report's intrinsic plausibility. They do not see widespread agreement among authorities as a natural sign that what they say is metaphysically true. Rather, they treat it as being commemorative in nature; formerly, where such agreement has obtained, the results of acting in accordance with it have tended to be beneficial; and so we are led to expect them to be so now (8 = Fr. 10b, 68–9 D). The Empiricists, by their own account, are subject to psychological compulsions,
end p.308

what Hume was to call the force of custom and habit. What marks them off from the laity and underwrites what claims they may have to expertise is simply the fact that their psychological habits have been formed by long practical engagement with the domain in question.

They have no theory regarding the patients' underlying physiology to explain the efficacy of these empirically derived cures: all such aetiologizing is both intrinsically undiscoverable and therapeutically irrelevant. All that matter for Empiricists are the appearances, the *phainomena*; and Empirical medicine consists solely in the careful accumulation of data concerning which types of observable phenomenon have tended to go with which, and what sorts of therapeutic intervention have been seen to help in which types of case. 'Experience', *empeiria*, or the collection of such regular conjunctions is described as

328 the memory of what one has seen to happen often and in the same way. (4 = Fr. 10b, 50 D)

The Empiricist doctor's basic requirement is simply to be able to recall how things have turned out in the past, and to use those memories to guide his therapy.[30]

(c) The Rationalists' Counter-Attack

Thus the Empiricist epistemology is built upon direct experience, and derivatively upon the reports of the direct experiences of others. It deliberately rejects theoretical reasoning about the underlying structures of things (if any) in virtue of which they exhibit what properties they do. If science is the business of supplying such aetiological explanations, then they are not in the business of science—but that fact, if it is one, will not bother them at all, provided that they can make good their claims to the pragmatic efficacy of their procedures, and provided that they can maintain their official position that medicine as they practise it (and indeed as it ought to be practised) makes no claims upon theory of any kind.

At this point, however, the Empiricists are vulnerable to Rationalist attack. Galen's *On Medical Experience* is couched in the form of a dialogue between an Empiricist and a Rationalist concerning method (the Rationalist arguments are largely drawn from Asclepiades of Bithynia (fl. *c.*125 bc),[31] the Empiricist defence being owed in large part to the second-century ad doctor Menodotus: § *f* below). First of all, the Empiricists'

end p.309

account of the building-up of an *empeiria* relies upon their being able to recognize similarities between different cases. But, as Galen's Dogmatist points out, each case differs from every other in some respects, while unrelated cases may share irrelevant features: 'is it your opinion, then,' he asks

329 that it is only in the subject with which physicians are concerned that the memory has simply to deal with simple and isolated things, and that there are no combinations and varieties in it? If you do affirm this, we would say: what is more manifold, more complicated, and more varied than disease? In any case, how does one discover that a disease is the same as another disease in all its characteristics? Is it by the number of symptoms or by their strength and power? For if a thing be itself, then, in my opinion, it must be itself in all these characteristics, for if even one of them is lacking, it is perverted and ceases to be itself, since it no longer possesses the quality lacking. (Galen, *On Medical Experience* 3–4, 89 Walzer, trans. after Frede and Walzer[32])

What is it about certain factors in the various cases that makes them similar? And how can this be established without relying on theory of some kind (ibid. 3–6, 88–93 Walzer)? Any taxonomy relies on some (at least implicit) theory of relevant similarity, which the Empiricist, eschewing all theory, cannot supply.

But even if this difficulty can be overcome in the case of personal observation, what is to be done with *historia*?

> **330** How can a person determine whether what he sees at this moment is identical with that which someone else has seen before or is something quite different, unless he himself has seen both? (4, 89 Walzer)

Lacking a theory of causal relevance, the Empiricist has no way of telling which antecedent factors are relevant to the individuation of the disease. Given that they are indefinitely many in number, he cannot enumerate them all; but he has no criterion with which to make a judicious selection among them (6, 91–2 Walzer).

But even if the Empiricists can sidestep these objections, a fatal difficulty lurks at the heart of their concept of an experience, 'that which has been observed many times': *how* many times? The Rationalist here constructs a sorites. One observation is insufficient; but if one is insufficient, so are two, and if two so too are three, and so on 'until I reach a very high number' (7, 96 Walzer).[33] If one observation cannot make an experience, then the addition of one to a previous set cannot make the difference either.

end p.310

As regards similarity, the Empiricist simply denies that he needs any criterion. He is not in the Rationalist business of speculating about what grounds the similarities; rather he simply acts upon apparent resemblances. An ailment observed today may look like one he saw last week; and he will perhaps be able to say why it appears so. Since he makes no attempt to *justify* his practice other than pragmatically, he has no need for any theory of relevance in resemblance.

In the same vein, the Empiricist neither has nor requires a single answer to the question how many times makes many (16–18, 114–20 Walzer): that varies from case to case, being a psychological rather than a logical issue. No inferences are involved in constructing an experience; the Empiricist simply comes to discern a general pattern in a set of particular cases. The pattern may turn out to be misleading: if so, experience will lead to its abandonment or modification. But experience itself suggests that it will not prove deceptive.

(d) Transition to the Similar

More controversial for the Empiricists was their account of the analogical procedure they called 'transition to the similar'. This was not conceived even by those Empiricists most inclined to its use as a method of discovery; rather it functioned as an empirically tested means of generating new testable hypotheses (*Outline of Empiricism* 9 = Fr. 10b, 71 D). 'Transition' occurs when one sees that a particular case is, while not actually a recurrence of some previously experienced type, at least sufficiently similar in structure to one which has been in potentially relevant ways. Suppose you have dislocated a hip, an injury

which I have never personally encountered, and about which *historia* is silent. Even so, if I have previously reset dislocated shoulders, I may reason that shoulders and hips are morphologically similar, and hence that what proves efficacious for one may well work for the other. This reasoning justifies me in trying the shoulder treatment on the hip.[34] Moreover, transition is not restricted to moving analogously from one part of the body to another homologous with it. One can employ transition more or less anywhere in the theorems—I might decide that some drug seems reasonably similar to one I have successfully used in the past, but run out of (Galen, *On Sects* 2, 3–4 H; cf. *Outline of Empiricism* 3, 9 = Fr. 10b, 49, 69–74 D). Indeed
end p.311

331 in the case of transference of one remedy from one ailment to another similar to it one has a greater or smaller basis for expectation of success in proportion to the increase or decrease in similarity of the ailment, whether or not *historia* is involved. And the same goes for the transference from one part of the body to another. (ibid. 9 = Fr. 10b, 74 D)

The greater the number of similarities between the tested and the proposed cases, the better founded is the transition: but that is not to make any metaphysical assumptions about the regularity of the universe of the type that Hume thought fatally compromised the notion of induction (*Enquiry concerning Human Understanding*, iv . ii). The Empiricist should not be greatly surprised if the results are in some cases disappointing, since he has no Rationalist pretension to have discovered the hidden reason why things behave as they do.

The Empiricists carefully differentiate their method from that of the Rationalists, which might at first sight seem to be the same:

332 Rationalist transition based on the nature of things lays hold of knowledge by means of indication (*endeixis*).[35] But the Empirical variety relies on what is discovered by experience,[36] not because it is persuasive or plausible that the similar should be productive of something similar, or require similar things, or undergo similar things; it is not on the basis of this, or anything else of this sort, that they think it justifiable to make the transition, but on the basis of the fact that they have discovered by experience that things behave this way. (Galen, *Outline of Empiricism* 9 = Fr. 10b, 70 D)

The practice of transition to the similar is itself justified on the grounds of past empirical success.

Thus the Empiricists are able to give an account of their own procedures which makes no reference to any real states of affairs beyond the appearances in virtue of which their practice works. If it does work, that is simply a brute fact, one which cannot, in view of the evident disagreement among the Rationalists, be explained in terms of its convergence to the real structure of the world.

(e) Signs and Inference

Transition to the similar is evidently a procedure of considerable epistemological potential, the more so the more things are taken to fall within its legitimate ambit. But equally, the more extensive its use, the more it
end p.312

apparently compromises the methodological purity of Empiricism. Empiricism is founded, after all, on the rejection of pure reason, by way of indicative signs, as a mechanism for arriving at the truth; and while the role of reason in Empirical medicine is itself controversial, it is evident that the more weight that is placed on transition, the harder it becomes to maintain that the Empirical method is one untainted by theory. Thus is it unsurprising to discover that whether or not transition was to be allowed, and if so in what circumstances, were questions hotly debated among the Empiricist doctors themselves:

333 the question has been raised whether Serapion also believes that transition to the similar is a third constitutive part of medicine as a whole. Menodotus thought that it was not, but that the Empiricist merely makes use of transition to the similar—and it is not the same thing to make use of something as to treat it as a part. Cassius the Pyrrhonian, furthermore, tried to show that the Empiricist does not even make use of transition of this sort; indeed he has written an entire book on the subject. Theodas [fl. second century ad] did better when he held that transition to the similar constituted reasonable experience. Yet others, though, have claimed that transition to the similar is more like an instrument. (Galen, *Outline of Empiricism* 4 = Fr. 10b, 49–50 D)

Despite the obscurity of some of that, there was evidently a spirited and complex dispute about transition's status and role. The disagreement about whether transition was a proper part of Empirical medicine or simply something Empiricists made use of resembles that between Stoics and Peripatetics on the role of logic: for the former it was one of the three parts of philosophy, along with physics and ethics (Diogenes, *Lives* 7. 39; Sextus, *Against the Professors* 7. 19), while the latter viewed it rather as a tool with which philosophy itself could be pursued (Ammonius, *On 'Prior Analytics'* 8–9). Perhaps transition is not on a par with personal experience and *historia* as a constitutive part of the Empirical method (transition only suggests procedures, which must then be subjected to further testing; it is part of the context of discovery, not that of justification: Galen, *Outline of Empiricism* 9 = Fr. 10b, 71 D). But it may also concern the extent to which one may have warranted confidence in the result of some application of the method. The more the Empiricist tends towards genuine scepticism, the less likely he will be to admit that transition is a proper part of Empiricist medicine (since it apparently involves going beyond mere appearances), although he may recognize himself using it; and if he arrives at the position of Cassius, he will not even do that.

Celsus describes Cassius as 'the greatest doctor of our times' (*On Medicine*, Proem 69 = Fr. 30 D).[37] His refusal to allow transition any role at all

end p.313

in Empirical medicine, even as an adjunct to it, was presumably part an attempt to purge Empirical practice from what he saw as unwarranted Rationalist accretions. And presumably this took the form of rejecting anything that smacked of reasoning and inference. For these tougher Empiricists, Empirical practice was simply practice. To allow inference into the picture at all would compromise their differences from the

Rationalists and to subvert their attack on reason and theory.[38] In line with this tough stance, some heroically held that even complex drugs could be discovered by chance alone, unsupplemented by any theory (Galen, *On the Therapeutic Method* 10. 163). On the other hand, even though he allows that some Empiricists are respectable practical physicians (*On Sects* 4, 7–9 H), Galen baulks at the possibility of complex remedies, or technical instruments such as the cupping-glass, being discovered by improvisation and sheer luck (*On the Affected Parts* 8. 154).

It is tempting to speculate that the Empiricists gradually relaxed their original tough stance in the face of Rationalist objections, at the same time and in the same manner as the gradual convergence of Stoic and Academic epistemology.[39] And just as Aenesidemus reinvented Pyrrhonism in response to what he saw as the unacceptable watering-down of Academic scepticism, so the irredentist Cassius swam against the syncretist tide that threatened to engulf the original, genuine Empiricism.

We need, then, to determine what precisely this hardline Empiricism consisted of, as well as investigate how revisionist Empiricists could distinguish themselves from the Rationalists. To make progress with both of these issues it will help to establish what, for the ancients, reason amounts to.[40] From Aristotle onwards (indeed perhaps earlier) the Greeks tended to conceptualize reason as a faculty by which the soul may progress beyond immediate experiential data to form general conclusions. But, sceptics will object, the gap between experience and rationally derived ‘knowledge’, which here amounts to an understanding of the causes of things,[41] is unbridgeable; and the Empiricists make use of the sceptical trope of undecidable dispute (cf. VIII.4) at precisely this juncture (Galen, *On Sects* 5, 11 H), that dispute being an indication of ‘the non-apprehensibility of things’ (cf. **335** below).

Moreover, the Empiricists attack the Rationalist concept of proof:[42]

end p.314

334 they do not grant that there is such a thing as indication (*endeixis*) or that one thing can be known on the basis of something else, for one has to know everything on the basis of itself. Nor do they allow that there is such a thing as a sign of something which by its very nature is non-evident. Furthermore, they argue that no art (*technē*) has any need of logic. . . . Then they talk about the fallacious modes of proof which the Dogmatists are accustomed to use and in particular about the class of analogisms. . . . Epilogism, on the other hand, which they describe as reasoning solely in terms of what is apparent, is of use in the discovery of things which are temporarily non-evident. For this is how they themselves call things which are by genus perceptible but which have not yet become apparent. (Galen, *On Sects* 5, 10–11 H)

Analogism (*analogismos*) is the Empiricists' equivalent to indicative sign-inference, inference to permanently hidden conditions; epilogism answers to the commemorative sign. Indication (*endeixis*) is the method by which the Rationalists sought to derive therapeutic information on the basis of pathological theory (cf. Galen, *On the Therapeutic Method* 10. 242–9): external signs serve to indicate the hidden conditions of the body.

But that is exactly what the Empiricists deny can be done. If such inferences really were secure, they would be undisputed; yet sceptical philosophers treated medical disagreements as paradigm cases:

335 in the case of fever patients, flushing and prominence of the vessels and a moist skin and increased temperature and quickening of the pulse and all the other signs of the same thing . . . nor do they appear alike to all; but to Herophilus, for example, they seem to be definite signs of good blood, to Erasistratus of the transference of the blood from the veins to the arteries, and to Asclepiades of the lodgement of theoretical particles in the theoretical interstices. (Sextus, *Against the Professors* 8. 219–20)

Aenesidemus used that instance of medical disagreement to argue that a universally accepted indicative sign is a chimera, while no uncontroversial criterion exists with which to press the claims of one theory over another, and without such a criterion disputes of this sort are simply undecidable (cf. Aenesidemus against the aetiologists: VIII.2); but it is probably owed to the Empiricists, who say that

336 inapprehensibility (*akatalēpsia*) is the cause of the undecidable dispute (*diaphōnia anepikritos*), while the dispute is in turn a sign of the inapprehensibility. And they note that it is the dispute concerning non-evident matters which cannot be decided, not the dispute concerning evident matters. For in the latter case everything, once it is apparent what it is like, confirms those who are right and refutes those who are wrong. (Galen, *On Sects* 5, 11–12 H)

These passages presumably represent the Empiricism of Menodotus and Theodas, Galen's older contemporaries. The Cassian hard line is absent: a

form of reasoning is to be tolerated. But it cannot lead to the establishment of claims about matters that are inevitably beyond the scope of empirical assessment, such as the internal states of a patient's body. And if the Rationalists counter that such information is available to the vivisectionist, the Empiricists will reply that what we need to discover is how those internal processes operate in their normal state: but no one can maintain that the condition of a creature undergoing vivisection is normal (Galen, *On Anatomical Procedures* 2. 281–3; Celsus, *On Medicine*, Proem 40–4; cf. 74). Anatomy is useless, and vivisection mere pointless cruelty.

Thus we have a provisional answer to a question canvassed earlier: reasoning is acceptable if it is at least susceptible of empirical confirmation. Epilogistic, commemorative sign-inferences are acceptable, as others are not, since their claims can be verified. That contention is not, of course unexceptionable, and is vulnerable to Humean attack. What are the criteria for verification, and how may one know that they are satisfied? Perhaps Cassius 'the Sceptic' asked precisely those questions.

Whatever the truth of that, the earliest Empiricists attempted to justify their stance in severely commonsensical terms. Knowledge simply amounts to the possession of a body of apparently reliable general beliefs induced by a sequence of perceptions stored in the memory,[43] and 'accessed' as a result of stimulation by perception of an appropriately similar condition. Now that description appears to presuppose some theory of behavioural psychology; but it is open to the Empiricists, as to Hume, to reformulate it as a non-dogmatic description of the facts of psychological life, a further fact of which will be that the doctor who proceeds along these memorist lines will often appear to have made the correct decision (in that his intervention will be followed by the patient's recovery). But what if the patient simply recovered spontaneously? Surely a theory of medical aetiology is needed to underwrite the claim that it was the doctor's action that made the difference? Not for the Empiricist: he will continue to act according to his (limited) beliefs: but they are not unalterable, and the Empiricist will continue to test their efficacy against further cases, modifying them as the need arises.

Finally, it is possible that some underlying cause really does explain the persistent success of some beliefs about phenomenal correlations—but it does not matter to the Empiricists whether there is or not. They make use of commemorative signs, but endorse no metaphysical explanations for their apparent success. Even in the case of epilogism, Empiricists can distinguish between endorsing a practice and simply employing it (cf. **333** above, on transition).

(f) The Empiricism of Menodotus

The hardline Empiricist, then, rejects all reasoning (or rather rejects it *as* reasoning). But texts like **333** betray the fact that the Empiricists disagreed among themselves about reason, a fact confirmed by Galen's *Outline of Empiricism*. In chapter 11 (Fr. 10b, 80–6 D) Galen, discussing 'the language appropriate to an Empiricist', castigates the school for engaging in vehement controversy, a practice supposedly unbecoming to those who profess the quietist ideals of an Empiricist life: they should let their deeds alone speak for them. Galen accuses Menodotus of resorting to invective in place of argument, and more seriously of crypto-Dogmatism: he had written a book designed to show that Asclepiades' views were entirely false,

337 in spite of the fact that on innumerable occasions he has said that one should approach everything non-evident as if it perhaps is, and perhaps is not true. (Galen, *Outline of Empiricism* 11 = Fr. 10b, 84 D)

Galen may not be playing quite fair here: Asclepiades had, after all, mounted the most sustained and devastating critique of Empiricist epistemology, and it would be neither inappropriate nor inconsistent with Empiricist practice to produce a counter-blast exposing the frailty of its arguments; such a practice need not commit Menodotus to actually endorsing his own arguments, and hence slipping into a species of negative dogmatism.

Yet Sextus, who is usually scrupulous about the importance of distinguishing genuine scepticism from negative dogmatism, sets Empiricism apart from Pyrrhonian scepticism on the grounds that it

338 positively affirms the non-apprehensibility of the non-evident, (*Outlines of Pyrrhonism* 1. 236)

and he probably had Menodotus mainly in mind. Whether or not Sextus is right to distinguish Empiricism from scepticism on the grounds of its allegedly negative dogmatist tendencies (cf. the claim of **333** that dispute is a cause of non-apprehensibility), Menodotus' Empiricism was clearly more accommodating than the hardliners' version:

339 Menodotus frequently introduces a third thing in addition to perception and memory, which he calls ‘epilogism’; sometimes, however, he does not posit anything in addition to memory except perception. (Galen, *Outline of Empiricism* 12 = Fr. 10b, 87–8 D)

This is obscure:[44] but the most plausible interpretation of it has Menodotus allow memoristic and epilogistic accounts of theorem-building to coexist, albeit with different statuses. Memorism was devised to show that the

end p.317

Rationalist insistence upon the necessity of reasoning is ill-founded; while actual Empirical practice invokes epilogism, in now-familiar fashion, without thereby endorsing it. Menodotus remains resolutely sceptical after all.

For all that, the dispute within the school about the role of reason was not merely an empty one: some Empiricists, most notably Heraclides of Tarentum (fl. *c.*75 bc : see Galen, *Outline of Empiricism* 12 = Fr. 10b, 87 D), were prepared to posit the existence and utility of a faculty of human reason. Moreover, we know from other sources that Heraclides was prepared on occasion to speculate as to why certain outcomes had indeed transpired: he argued, against the prevailing Rationalist orthodoxy, that on occasion luxations of the hip could be successfully reset (n. 34 above); and in those cases the tendon which holds the hip-bone in place is not severed, as the Rationalists claim it invariably is when dislocated, but merely stretched (Galen, *On Hippocrates' ‘On Joints’* 18a 735 = Fr. 175 D).[45] Even so, it is not clear that Heraclides actually adopted this explanation for himself, as opposed to merely using it to demonstrate the fallibility of a particular piece of Rationalist reasoning. This case, then, in no way implies that even the most accommodating Empiricist would actually endorse anything like indicative sign-inference.

3. Methodism

(a) The Structure of Medical Methodism

The origins of the Methodist school, which arose early in the first century ad , are obscure;[46] but its doctrines were striking and innovative. Contrary to orthodoxy, both Empiricist and Rationalist, they held that medicine was a relatively simple matter, and could be learned in a matter of months. They disregarded the patient's causal history as therapeutically irrelevant,

340 claiming that the indication as to what is beneficial, derived directly from the affections themselves, is enough for them, and not even these taken as specific particulars, but taking them to be common and universal. Thus they also call these affections which pervade all particulars ‘communalities’ . . . which they call constriction and relaxation, and they say that each disease is either constricted, relaxed, or a mixture of the two. (Galen, *On Sects* 6, 12–13 H)

end p.318

Only two basic pathological conditions exist: relaxation, where ‘the bodily fluids flow too freely’, and constriction, where their normal flow is impeded; although some conditions involve a mixture of both types. The doctor's only concern is to determine which of these states the body is in, a simple task (they claim) for anyone with a modicum of practice.

Sextus in fact reckons the Methodists to be more Pyrrhonian than the Empiricists, since

341 the Methodist speaks of ‘communality’ and ‘pervade’ and the like in a noncommittal way. Thus also he uses the term ‘indication’ undogmatically to denote the guidance derived from the apparent affections or symptoms, both natural and unnatural, for the discovery of the apparently appropriate remedies. (*Outlines of Pyrrhonism* 1. 240)

And the Methodist is driven by the ‘compulsion of the affections’, as the Sceptic is, to treating conditions with contrary remedies. Methodism, then, does not, apparently, even involve memory. Far from having to gain an understanding of the communalities as a result of long experience, they are, it seems, straightforwardly accessible to direct perception. If Empiricism is metaphysically parsimonious, Methodism is positively miserly.[47]

Equally, the official Methodist doctrine, while allowing signs, has no truck with sign-inferences as such. The fifth-century ad medical writer Caelius Aurelianus[48] reproduces an early Methodist argument:

342 Thessalus and his sect . . . argue thus: 'if there were sure and inevitable signs of future events, such as the onset of phrenitis, all who manifested them would necessarily develop phrenitis. But some of those who show these symptoms do not develop phrenitis. (Caelius Aurelianus, *On Acute Diseases* 1. 22)

Moreover,

343 every sign is understood in relation to what is signified, since signs belong in the category of relations. But can anything be called a sign if the thing signified is not only not present now, but in some cases never will be? (ibid. 1. 29)

All the same, the Methodists allow that there is such a thing as indication, although not in the Rationalist manner; Galen remarks that

344 however much they occupy themselves with what is apparent, they are distinguished from the Empiricists by their use of indication (*endeixis*). . . . And the Empiricists, they say, will have nothing to do with anything non-evident,

end p.319

claiming it is unknowable, while they themselves will have nothing to do with anything non-evident since it is useless. Furthermore, the Empiricists derive observation from the apparent, while they derive indication from it. (Galen, *On Sects* 6, 14 H)

The ‘indications’ of the Methodists are of the relaxed and constricted states, something which at first sight seems to place them closer to the Rationalist camp; but they insisted, in a way that Galen for one found dubiously coherent, that such indications require no complex theory, and are in a sense matters of direct observation.

(b) The Dispute About Causes

The most important source of difference between the three major schools in the early Imperial period, however, concerned causes. The Empiricists, even though they eschewed causal theory as such, happily talked of antecedent causes (Galen, *On Sects* 4, 7 H), and made room for them in their general account of the features relevant to a particular case, the *sundromē*. Take a case of rabies: the Empiricist will treat as relevant, indeed crucial, to the case that the fatal bite was that of a mad dog (7, 18 H), since such cases have been observed to turn out catastrophically differently from ordinary bites. By contrast, Galen says, Methodists allow no therapeutic relevance to the state of the dog (it forms no part of the indication); they will treat this bite no differently from any other (8, 19 H). While Rationalists, although accepting the relevance of the dog's mental condition, will not simply rely on experience but will try to explain theoretically how the dog's insanity can bring about hydrophobia.

Galen affirms that the Empiricists 'do not hesitate to ask for the so-called antecedent cause' (4, 8 H), a claim apparently confirmed by their attitude to rabies cases. On the other hand Galen says elsewhere that the Empiricists 'doubt whether there are causes or not' (*On Antecedent Causes* 13. 162), remarking a little later that

> **345** even those doctors from the Empirical school, who above all others proclaim things in accordance with common sense, were so overcome by the sophism as to be moved to doubt concerning antecedent causes. (ibid. 13. 170)

The 'sophism' is Erasistratus' denial of causal status to antecedent causes (§1*c*, **322–5**; XI.2*a*).

This apparent inconsistency may be defused by a proper understanding of the sense in which an Empiricist accepts antecedent causes. No causal theory is involved: he essays no account of how the dog's bite can produce

end p.320

rabies. Indeed, he does not, strictly speaking, really know that it does. He simply recognizes that rabies supervenes upon bites of that type. What the Rationalists will treat as being a genuine cause (which as such demands that some explanation of its operation be true), the Empiricists rather take as a mere sign, an event suitably concatenated with some future outcome. Call them causes if you wish (Empiricists, like Pyrrhonists, do not argue about terminology), as long as the term is not taken to connote any commitment to the existence of some arcane causal truth of the matter. Empiricists and Rationalists, then, largely agree (against the Methodists) about the importance of those events classed by the Rationalists as antecedent causes; they disagree about what can be said as to why they are important. The Empiricist allows antecedent causes into his syndrome; but he does not thereby commit himself to believing that they really are *causes*.

In the light of this it is less surprising that the Rationalist Galen can allow that some Empiricists are perfectly respectable practitioners. In fact, he thinks he can explain it:

346 the same things from which the Dogmatists derive the indication of what is beneficial form the basis of the Empiricist's observation. For the collection of symptoms in the case of the person who has a fever, which they are accustomed to call the ‘syndrome’, suggests evacuation to a Dogmatist, but to an Empiricist the recollection of his observation. . . . And in general the Dogmatists and the Empiricists draw on the same medicines for the same affections. What they disagree about is the way these remedies are discovered. For, given the same apparent bodily symptoms, the Dogmatists derive from them an indication of the cause, and on the basis of this cause they find a treatment, whereas the Empiricists are reminded by them of what has happened often in the same way. (Galen, *On Sects* 4, 7 H)

The Rationalist (or Dogmatist) thinks that the ‘nature of the matter’ (i.e. things' physical structures and how they interact) explains how and why the therapies work as they do. In so far as the Empiricist offers explanations at all, he does so only epistemically, with no metaphysical commitment whatsoever. He can give reasons for his adoption of a certain practice, and what gives him his limited and intrinsically defeasible confidence in it; but he has nothing to say about the nature of the reality, if there is one, which accounts for the success of his procedures.

4. Conclusions

Doctors worried about explanation. Diocles thought it a mistake to try and give causal explanations for everything (§1*a*); Herophilus exalted the appearances, and held that causal theorizing was at best provisional, and

end p.321

while he clearly indulged in some causal speculations himself, none the less produced a sceptical argument against the possibility of causes (§1*b*); while Erasistratus insisted that only proximate, sufficient causes were properly causes at all (§1*c*). More thoroughgoing in their scepticism were the Empiricists, who insisted that no useful knowledge of internal conditions was to be won, and that observation and report of evident conditions and their cures sufficed to construct medical science (§§2*a*–*b*). This position invited a counter-attack from more theoretically inclined doctors, who held that the Empiricists' own methodology was inadequate: they had no account either of the proper size of an inductive sample or of what sorts of similarities between cases were relevant: such things, they alleged, require richer theoretical resources than those available to Empiricism (§2*c*). Some Empiricists were, however, disposed to allow analogical reasoning a limited role in medicine, while still denying that it provided a route to determining the hidden conditions (§§2*d*–*f*). A third school, the Methodists, insisted that direct observation was enough on its own to indicate the appropriate treatment §3*a*, and refused to allow evident antecedent factors even the semiotic role accorded them by the Empiricists (§3*b*). Thus practical science provides the backdrop to a lively theoretical debate about the nature of causal explanation, its availability, and its usefulness.

end p.322

X The Age of Synthesis

R. J. Hankinson

1. The Origins of Syncretism

(a) The Academy of Speusippus

In tracing the concepts of cause and explanation through the Hellenistic period, we have dealt with the contributions of the Epicureans, the Stoics, and the medical schools, as well as that of Theophrastus and Strato in the Lyceum; the later history of Aristotelianism is shrouded in obscurity, at least until the first century bc , and it is not until the end of the second century ad that a figure of genuine stature re-emerges in Alexander of Aphrodisias (§4 below). One major school of the period has been largely disregarded, namely Plato's Academy, principally because after its sceptical turn under the scholarchy of Arcesilaus (*c*.272 bc) it evinced little interest in questions of cause and explanation, devoting itself principally to a subtle dialectical refutation and subversion of the Stoics' optimistic epistemology[1] (an exception being Carneades' assault on divination and the doctrine of fate: VI.3*a*; VII.2*d*; VIII.5*b*).

However, Platonism underwent a series of changes in the years between Plato's death and Arcesilaus' sceptical *démarche*, under the leadership first of Speusippus (*c*.407–339 bc), Plato's nephew and intellectual legatee, and then of Xenocrates (396–314 bc). Although it may seem somewhat untidy to treat of them here, their importance for our purposes lies largely in the influence their modification of Platonic doctrine had upon the Platonists of a later, more eclectic age: those Platonists, that is, who sought to rehabilitate a doctrinal Platonism after the collapse of the sceptical Academy in the early part of the first century bc .[2]

Nothing of Speusippus' work survives complete; but it is evident from his fragmentary remains that he sought to push Platonism towards an even closer *rapprochement* with Pythagoreanism. His influence over the Platonism of his immediate successors was minimal: but in many ways his views laid the foundation upon which Plotinus and the Neoplatonists were

end p.323

to build (XII.2*a*).[3] He adopted as fundamental the doctrine of the distinction between Unity and Plurality (in the form of the Limit and the Unlimited) to be found in *Philebus* 16c–18b (cf. III.4*d*), which indeed has Pythagorean precursors in the views of Philolaus (I.4*a*, **46–8**). His One, like Plato's Good (*Republic* 6. 509b), was, in the phrase that was to become a staple for the Neoplatonists, ‘beyond being’:

347 such is Speusippus' interpretation of the views of the ancients. What does he say? 'Holding the One to be beyond Being, and the source from which Being derives, they freed it even from the status of a principle. But considering that if one took the One in itself, conceived as separate and single, without anything else, and adding no other element to it, nothing else would come into being, they introduced the Indefinite Dyad as the principle of all things.' (Proclus, *On the 'Parmenides'* 38. 32–40.5K = Speusippus, Fr. 62 IP[4])

Thus the Indefinite Dyad is required in order for there to be any generation of anything at all. He did not stop there:

348 Speusippus posited many substances beginning from the One, and principles for each type of substance, one for numbers, another for magnitudes, another for the soul—in this manner he multiplies substances. (Aristotle, *Metaphysics* 7. 2. 1028^{b}22–5 = Fr. 48 IP)

The One is not only beyond being, but beyond goodness as well, a view Aristotle castigates:

349 those who believe, like the Pythagoreans and Speusippus, that the most beautiful and the best do not exist in the beginning, on the grounds that while the origins of plants and animals are their causes, the beautiful and the complete exist in what comes from them [i.e. from the origins], are mistaken. (12. 7. 1072^{b}30–5 = Fr. 53 IP)

Speusippus' One first generates the arithmetical numbers, and then, in obscurely neo-Pythagorean fashion, has them produce the furniture of the physical world out of receptive matter:

350 the elements from which the numbers derive are themselves neither yet beautiful nor good; but from the union of the One and the cause of multiplicity Number arises, and it is here that Being and Beauty first appear. Next in order there arose from the elements of lines the geometrical realm, in which there is also Being and Beauty, in neither of which is there anything ugly or evil. It is at the lowest level, in the fourth and fifth realms, which are put together from the lowest elements that Evil arises, and even then not in a primary fashion, but because of the failure to control some part of the substrate. (Iamblichus, *On the General Mathematical Science* 4. 18. 1–12 Festa = Fr. 72 IP; cf. XII.3*a*, **469**)[5]

end p.324

That account of the origin of evil anticipates later Platonist and Christian accounts of the same troubling phenomenon. Speusippus evidently tried to describe a progression of creation away from the pure realm of the One down to the lowest material level. Aristotle objected that Speusippus, in trying to protect the transcendental unity and separateness of the One, is forced to compromise the metaphysical harmony of the universe as a whole:

351 as for those who hold that mathematical number is primary, and so are always producing another substance and other principles for each of them, they make the substance of the whole episodic[6] (for each thing contributes nothing to any other, either by being or not being), and many ruling principles. But reality should not be badly ruled. (Aristotle, *Metaphysics* 12. 10. 1075^{b}37–1076^{a}4 = Fr. 52 IP)

Speusippus rejects orthodox Platonic ideas, preferring, in Pythagorean fashion, to elevate numbers to the role of principles. Nor is the One identical with God, who is, however, the same as Intellect (Fr. 89 IP): hence the One is distinct from intellect as well. The fragmentary state of our evidence for Speusippus' metaphysical doctrine makes it impossible to determine the extent to which Aristotle's typically trenchant criticisms are justified; but while the details of the progression from the One to the other levels of reality are, at least in our sources, obscure, and perhaps merit Aristotle's criticism of it as episodic, we may also here descry the remote intellectual ancestor of the grand metaphysics of the Neoplatonists.

(b) Xenocrates

Speusippus' successor Xenocrates had a more immediate effect on the Platonic tradition. He too adopted the fundamental metaphysical division between the One and the Dyad, but in other ways (notably in his identification of God and Intellect) he parts company with Speusippus:

352 Xenocrates . . . said that the Monad and the Dyad are gods, the former, being male, having the role of the father ruling in the heaven, which we call Zeus, odd, and intellect, who is for him the primary god. The latter, being female, as mother of the gods rules things under the heaven, and is for him the soul of everything. The heaven is a god, and the fiery stars are Olympian gods, and there are other unseen sublunary demons. He held the corporeal elements contained .[7] (Aëtius 1. 7. 30 = Fr. 213 IP)[8]

end p.325

Moreover, he apparently made a more concerted effort than his predecessor (cf. **351**) to show how the whole of reality was generated from the principles. A text from Theophrastus' *Metaphysics* bears quoting:

353 now most people come to a halt [sc. in metaphysical explanation on the basis of principles] having reached a certain point, as do those who establish the One and the Indefinite Dyad: for, having generated numbers and planes and bodies, they pass over practically everything else, touching on them merely, and clarifying only this, namely that some things (such as place, void, and the infinite) come from the Indefinite Dyad, while others come from numbers and the One (such as soul, and certain other things); time, and the heaven, and some other things come to be at the same time, but they say nothing more of the heaven and all the rest. Speusippus and his associates do this, and everyone else, except for Xenocrates, who does in a way give everything in the universe equally a place, whether it be perceptible, intelligible, mathematical, or divine. (Theophrastus, *Metaphysics* 2. 11–13 = Fr. 100 IP [= Speusippus Fr. 87 IP])

Whether Theophrastus was right to think Xenocrates' derivation of reality less episodic is impossible to tell, since we cannot reconstruct in any detail the attempt Xenocrates made to move from the purely mathematical and geometrical primary reality to the structure of the physical world (a problem that beset Platonist cosmology from the *Timaeus* onwards).[9]

Speusippus had rejected Plato's Ideal Numbers; Xenocrates rehabilitated them with a vengeance. If Aristotle's remarks at *Metaphysics* 7. 2. 1028^{b}24–6 (Fr. 103 IP) refer to Xenocrates (as was thought in antiquity) he put Number and the Forms on the same metaphysical level, perhaps even identifying them. But the Forms are none the less responsible (at some removes) for the existence of particulars, although there are no Forms of particulars as such. A passage from Proclus' commentary on *Parmenides* bears quoting in regard to Xenocrates' views on cause and explanation (which were central to his philosophy: he defined wisdom as ‘the knowledge of first causes and intelligible substance’: Fr. 259 IP), even though it employs a vocabulary characteristic of a much later age:

354 The Forms are both intelligible and essentially immovable . . . and form- producing causes. By ascending to these principles they [sc. Platonists] make the whole of generation depend upon them, as Xenocrates says when he makes the Form the paradigmatic cause of things which exist naturally and for ever. Nor can it be placed along with the co-operative causes (*sunaitia*)[10] (by

end p.326

which I mean, for example, the instrumental, the material, and the formal), since it is completely the cause. Nor is it one of the final or efficient causes (at least in an unqualified sense). For although we say that it [i.e the Form] acts in its very being and is the end for things which are generated by assimilating them to itself, that which is genuinely the final cause of everything and that for the sake of which everything exists is prior to the Forms, while that which is genuinely the efficient cause is posterior to the Forms, looking towards the paradigm as criterion and measure; for being intermediate between the two, it desires the one, and is the object of the other's desire. So if it is the paradigmatic cause of things which exist naturally, there will be a Form neither of unnatural nor of artificial things; and if of things which exist for ever, there will be a Form of none of the things which are generated and destroyed serially. Therefore Xenocrates constructed this definition of Form as though it were in agreement with the Master [i.e. Plato], setting it up as a separated and divine cause. (Proclus, *On the ‘Parmenides’* 74 = Fr. 94 IP)

That passage requires little comment. But it is worth noting how a later writer at least interpreted Xenocrates as differentiating the transcendent Form from the role played by Aristotle's formal cause. This is probably an anachronism; but such a division of labour between the transcendent and the immanent formal principles (as well as the elevation of final and efficient causes above the others: see XI.3*b*; XII.3*d*) was to become a central feature of Middle (and Neo-) Platonism (cf. §2*d*, **365**; §3*d*, **378** below; XII.3*d*, **486**).

In physics, he made the rarified and the dense fundamental:

355 Xenocrates says that the stars and the sun are composed of fire and the primary density, the moon from the secondary density and its proper air, the earth from water and the third of the densities; in general neither density in and of itself nor rarefaction is receptive of soul. (Plutarch, *On the Face in Moon* 943f–944a = Fr. 161 IP)

This is obscure (and appears to conflict with other evidence suggesting that Xenocrates accepted Aristotle's ether: Frs. 264–6 IP; cf. V.2*f*); but the idea seems to be that the pure elemental constituents (presumably Plato's geometrical forms from the *Timaeus*: III.4*d*), which perhaps are to be identified with the rarefied, must be mixed with something more substantial to give them material reality; and that only such substantial composites are receptive of soul.

Soul itself Xenocrates famously defined as a ‘self-moving number’ (Frs. 165–88 IP, etc.) occupying a position intermediate between the purely transcendent realm of the Forms and the sensible world of material objects (Frs. 189–94). This doctrine excited Aristotle's

contempt as being ‘by far the most absurd’ of all the ones he had come across, his distaste for Pythagorean number-mysticism being much in evidence in the passage
end p.327

(*On the Soul* 1. 4. 408^b32–409^a10 = Fr. 195 IP); but, if obscure, it was without question influential.

Finally, a word on Xenocrates' demonology. Plato had already posited the existence of beings intermediate between gods and men (*Symposium* 202e; cf. *Epinomis* 984b–985b); Xenocrates apparently took over and fleshed out the notion (Frs. 22–30 IP), making some of these ethereal figures good and others bad; this theory, albeit in several different forms, was to become a commonplace of later Platonist doctrine (§3*c* below).

2. The Convergence of the Schools

(a) Posidonius of Apamea

Posidonius (*c.*135–*c.*50 bc), who was one of Cicero's philosophical mentors, and whose acumen in psychology and logic were much admired by Galen (cf. XI.5*c*), was a Stoic, and in some respects at least an orthodox one. He accepted the existence of divinatory signs of both a natural and a technical nature on the standard Stoic grounds of the universal sympathy and providential structure of the universe (Diogenes, *Lives* 7. 149 = Fr. 7 EK; cf. Frs. 106–8 EK): Cicero accuses him of excessive credulity, and perhaps even invention, on this score (*On Fate* 5–7 =Fr. 104 EK; cf. VIII.5*a*, **305**). His account of the cosmos as a single, living, intelligent entity (Diogenes, *Lives* 7. 142–3 = Fr. 99a EK) is equally orthodox, and he also probably upheld the Stoic doctrine of the periodic resolution of the world into fire (ibid. 7. 142 = Fr. 13 EK; cf. Fr. 97 EK), while other Stoics of the time, notably Panaetius, Posidonius' teacher, abandoned it in favour of the world's eternity (Philo, *On the Eternity of the World* 2; for Panaetius' rejection of astrology, see VIII.5*a*). His definition of causation equally toed the traditional line:

356 Posidonius says as follows: a cause is of something, because of which that thing is, either the primary producer or the instigator of the production; and the cause exists and is a body, while that of which it is the cause neither exists nor is a body, but is a property and a predicate. (Stobaeus, *Anthology* 1. 13. 1c [follows **256**] = Fr. 95 EK; cf. VII.1*c*, **257**)

On the other hand, Posidonius refused to accept that the extramundane void was infinite in extent (Stobaeus 1. 18. 4b = Fr. 97 EK), saying that it was only as large as was required in order to accommodate the periodic cosmic dissolution into pure fire; and he adopted the Platonic heresy that the soul was divided into three distinct faculties (for which Galen commends him: *On the Doctrines of Hippocrates and Plato* 5. 454–5, 480–1, 515 = Frs. 142–6 EK). Indeed, his psychology was, by comparison with
end p.328

the orthodox Stoic view of the strictly unitary nature of the soul, suspiciously dualistic: human beings contain both a rational and an irrational nature, and the latter is, if not properly schooled (ibid. 5. 432–5, 466–8 = Frs. 163, 31 EK), responsible for the existence of evil, which is thus natural (Galen, *That the Powers of the Soul Depend upon the Temperament of the Body* 4. 819–20 = Fr. 35 EK).[11] Furthermore, we know from

Galen's extensive report on the issue that Posidonius condemned as being explanatorily inadequate Chrysippus' view of the emotions as simply excessive judgements (*On the Doctrines of Hippocrates and Plato* 5. 424, 430–1, 460–1, 474–4, 502 = Frs. 156–62 EK).

Equally, his account of the physical world emphasizes, in a way foreign to earlier Stoicism, the separation between the active and the passive elements, between the cosmic intelligence and the matter within which it is manifest, even though he does not go so far as to treat them as entirely separable:

357 Posidonius said that there is a substance of all things and a matter without quality or form, in so far as it is assigned no particular shape or quality in itself, although it always has some shape or quality or other; and he says that substance is differentiated from matter, in that it has real existence, but in thought only. (Stobaeus 1. 11. 5c = Fr. 92 EK)

That opaque report has been variously interpreted;[12] but at the very least Posidonius wished to underline the conceptual separability of the active from the passive principles: the active is what gives shape and form to the receptive matter (ibid. 1. 1. 29b = Fr. 101 EK). Moreover, although he accepted that God pervaded the entire cosmos (Fr. 100 EK), he none the less stressed that the heaven (*ouranos*, presumably the outermost celestial circumference: cf. Aristotle, *Physics* 8. 10. 267^{a}20–^{b}9, on the location of the Prime Mover: V.2*f*) was to be identified with the ruling principle of the world (Diogenes, *Lives* 7. 139 = Fr. 23 EK). This view was neither unique nor original to Posidonius, but elsewhere (contrary to Stoic orthodoxy) he stresses the physical reality of limits:

358 surface is the limit of a solid body, having length and breadth without depth: Posidonius, in the fifth book of his *Meteorology*, maintains that it exists not only conceptually but in reality; (ibid. 7. 135 = Fr. 16 EK = 50e LS)

and such an emphasis is consistent with his holding that in some relatively strong sense the periphery of bodies is what determines what they are: 'Posidonius supposed that shape was the cause of differentiation and limitation and enclosure' (Fr. 196 EK); containing causes indeed. Thus the
end p.329

ruling principle, located at the periphery, supplies form to the matter of the world.[13]

Most significantly in this context, while orthodox Stoicism sought to identify Zeus with Fate (as well as with Reason: ibid. 7. 135), Posidonius is careful to distinguish them:

359 Posidonius said that Fate was third in order after Zeus; for Zeus is first of all, second is Nature, and third is Fate. (Stobaeus 1. 5. 15 = Fr. 103 EK)

These components of reality are not merely facets of one and the same thing, but rather stand in a hierarchical relation to one another.

In one other respect important for our purposes Posidonius goes beyond earlier Stoicism. As we saw in another context (VIII.2), Strabo remarked that

360 there is in him an abundance of aetiologizing and Aristotelizing, which our school avoids because of the obscurity of the causes. (Strabo 2. 3. 8 = T85 EK = **286**; cf. Seneca, *Letters* 95. 65 = Fr. 176 EK)

'Our school' is the Stoa (cf. Strabo 7. 3. 4): Strabo (64 bc –*c.*ad 22) clearly considered that Posidonius' predilection for causal explanation marked him out from the mainstream of the school.[14] That he was so predisposed is clearly borne out by the surviving evidence: **360** immediately follows a long discussion of Posidonius' view of the geographical and climatic structure of the world which is distinctly Peripatetic in flavour (ibid. 2. 2. 1–3. 8 = Fr. 49 EK). Moreover, he wrote, unusually for a Stoic, on the Aristotelian subject of meteorology (**358**; cf. V.2*d*, **191–6**; V.3*a*, **205–6**), dealing with such subjects as mock suns (Fr. 121 EK), comets (Frs. 131–2 EK), haloes (Fr. 134 EK), the rainbow (Frs. 15, 135 EK), thunder and lightning (Fr. 136 EK), earthquakes (Frs. 12, 23–3), snow (Fr. 11 EK), hail (Fr. 136), and winds (Frs. 137–8 EK).

Posidonius was also greatly given to empirical investigation. He devoted much time and trouble to investigating the phenomena of the tides (Frs. 214–20), criticizing Aristotle's account on the basis of his own detailed observations on the Atlantic seaboard of Spain, and recognizing the causal influence of the sun as well as the moon upon them (cf. VIII.5*a*; XI.1*b*). Such a notion is of course consistent with the Stoic concept of a universal cosmic *sumpatheia* mediated by the all-pervasive *pneuma* (VII.1*b*); the problem of action at a distance, which was to lead Galileo to reject with scorn the lunar theory of the tides, was not an issue for them. Nor did the question of tides exhaust his hydrological interests: he investigated the flow of rivers in various locations (Frs. 223–5 EK), and offered (as so many

end p.330

others had done) a theory of the flooding of the Nile (Fr. 222 EK), as well as concerning himself with questions relating to the relative levels of the land and the sea (Frs. 226–8 EK).

Finally, it has recently been argued persuasively (Kidd 1989) that Posidonius' monumental *History* was also designed fundamentally as an explanatory enterprise, designed to illustrate the effects upon events of the different national and ethnic characteristics, established on the basis of a detailed and comprehensive ethnography. One example, again reported in Strabo, will suffice: it was usual, apparently, to account for the migration of the Cimbri from Jutland (and the ensuing Cimbrian wars) on the grounds that they were forced from their homes by widespread flooding; Posidonius demurred, preferring to place the blame on their rapacious character: the inundation may have spurred them to action, as an antecedent cause, but it is not the real explanation for what followed (Fr. 272 EK; cf. Fr. 49 EK).[15]

(b) The Role of Astronomy

Posidonius busied himself too with astronomical matters, such as the description and explanation of various eclipse phenomena (Frs. 123–6 EK), as well as the determination of the sizes and distances of the sun (Frs. 19, 114–16, 120 EK) and the moon (Fr. 122 EK); furthermore, he produced one of the most accurate estimates thus far of the circumference of the earth (Fr. 202 EK). These latter concerns were commonplace among the Hellenistic astronomers. Aristarchus of Samos (*c.*310–230 bc), a pupil of Strato (Aëtius 1. 15. 5; cf. V.4*b*) and the inventor of the heliocentric theory of the solar system (Plutarch, *On the Face in the Moon* 6. 922f–923a; *Platonic Questions* 1006c: **388–9**: XI.

1*a*; cf. Archimedes, *Sand-Reckoner* 1), wrote a volume, which survives, computing *The Sizes and Distances of the Sun and Moon* on the basis of certain assumptions.[16]
Crucial among these is that sun and moon each subtend 2° of visible arc; that 'when the moon appears to us halved, the great circle dividing the bright and dark sections is in the direction of our eye': i.e. at half-moon, earth, sun, and moon form a right-angled triangle with the earth–sun side
end p.331

as the hypotenuse. Then, if the distance of the moon is known (Aristarchus estimates this on the basis of the ratio between its apparent diameter and that of the earth's shadow in lunar eclipse, plus an estimate of the earth's size), and the angle at the centre of the earth can be measured (by astronomical observation), a value can be determined trigonometrically for the distance of the sun, and hence (since the angle of arc which it subtends on the visual field can also be measured[17]) its size can be computed as well.
Aristarchus' actual results, particularly in regard to the sun's distance and size, are not very accurate: he estimated the angle subtended by the sun and moon at the earth's centre to be 87°, where the real value is 89° 50′, and thus inferred that the sun was only about 20 times more distant than the moon, and only 7 times larger than the earth, both of which underestimate the true values by a factor of around 20. But whatever its practical shortcomings, Aristarchus' attempt at cosmic calculation represents an unprecedentedly rigorous application of geometrical theory to a problem of celestial measurement.[18]
At about the same time Eratosthenes produced a relatively accurate measurement of the earth's circumference, again employing geometrical methods; he measured the angle off perpendicular of the sun's rays at Alexandria at the same time as the sun was directly overhead at Syene, some 5,000 stades (roughly 500 miles, but see n. **19**) further south, which he took to be on the same meridian; on the assumption that the sun's rays fall parallel to the earth's surface, he was able to deduce the earth's circumference to be 252,000 stades.[19] Eratosthenes' method is reported in Cleomedes'*On the Heavenly Bodies*[20] 1. 7;[21] the same chapter (Fr. 202 EK)
end p.332

also retails the observationally distinct (although geometrically equivalent) method of Posidonius: the bright star Canopus only becomes visible on the southern horizon (at maximum elevation), as one travels south, at Rhodes; however, at Alexandria, 5,000 stades south (and again supposedly on the same meridian) it reaches an altitude of 7.5°; yielding a value for the earth's circumference of 240,000 stades.
Thus Posidonius was actively involved in astronomical geometry. From our point of view, however, most significant is a passage on methodology, preserved from an epitome by the astronomer Geminus of Posidonius' *Meteorology*, copied by Alexander of Aphrodisias, and then by Simplicius in his commentary on Aristotle's *Physics*:[22]

361 physical theory seeks to consider the substance of the heaven and the stars, their capacity and quality, generation and destruction; and, by god, it can make demonstrations about their sizes, shapes, and positions. But astronomy attempts to say nothing of this sort, but demonstrates the arrangement of the heavenly bodies on the assumption that the heaven is a real cosmos, and speaks of the shapes, sizes, and distances of the earth, sun, and moon, and about eclipses and conjunctions of stars, and the qualities and quantities of their motions. Hence since it is linked to the consideration of quantity and size and quality in shape, it is reasonable that it require arithmetic and geometry; and those things of which it alone [i.e. astronomy] promises to give an account are established via arithmetic and geometry. Now the astronomer and the physicist will frequently set out to demonstrate the same thing, for example the size of the sun or the sphericity of the earth, but they will not do so by the same means. The latter will demonstrate each thing by way of essence, or capacity, or the fact that things are better thus, or generation and change, while the former will do so by way of the properties of shapes and sizes, and of the amount of motion and its relation to time. Furthermore, the physicist will frequently arrive at the explanation by looking towards the productive capacity, while the astronomer, demonstrating on the basis of external properties, is not able to provide an explanation, as for example when he shows that the earth or the stars are spherical; and sometimes he does not even desire to grasp the explanation, as when he discusses eclipses. And sometimes he seeks to offer certain models in a hypothetical manner, which are such as to account for the appearances if they are true. For instance, why do the sun, moon, and planets seem to move irregularly? If we postulate eccentric orbits for them, or that the heavenly bodies revolve on epicycles, their apparent irregularity will be accounted for; but it will be necessary to explain further what models could produce these appearances to make our study of the planets approximate to a causal account in accordance with a possible model. For this reason a certain Heraclides (of Pontus) has come forward and says that it is possible to account for the apparent irregularity of the sun if the earth

end p.333

moves in a certain way and the sun stays still in a certain way.[23] For it is no part of the astronomer's business to know what is by nature at rest and what kind of motions there are, but, by hypothesizing that some rest while others move, he investigates which hypotheses accord best with the celestial appearances. He should take his principles from the physicist, that the motions of the stars are simple, regular, and orderly, on the basis of which he will demonstrate that the dance[24] of all of them is circular, some of them spinning in parallel and others in oblique circles. (Geminus, in Simplicius, *On the ‘Physics’* 2. 2. 291. 21–292. 29 = Fr. 18 EK)

That passage may be compared with the following from Diogenes' survey of Stoic philosophy:

362 they say that cosmology is divided into two parts; in one inquiry the mathematicians too collaborate, the one where they investigate the fixed stars and planets, for instance whether the sun is the same size as it appears,[25] and similarly with the moon, and concerning their revolutions and similar inquiries. But there is another inquiry which is the province of the physicists alone, in which is investigated the substance of the cosmos, and whether it is generated or not, ensouled or not, and destructible or not, and if it is governed by providence, and so on. (Diogenes, *Lives* 7. 132–3)

Neither **361** nor **362** assert that mathematical astronomy is a purely instrumental science, although they differ slightly in what they take to be its proper scope. **362** reserves to astronomers the inquiry 'concerning the revolutions' of the heavenly bodies, and does not suggest that such an inquiry consists simply in elaborating more or less accurate predictive models; what is left to physics is the investigation of their substance. Equally, Posidonius in **361** holds that we must go beyond the mere practice of astronomical model-building if we are to supply genuine causal explanations for the phenomena which the models represent, and that cannot be done without the help of the physicist. The example of Heraclides is presumably intended to show that there can be mathematically equivalent models which none the less make radically different assumptions about what is in motion and what at rest: distinguishing between them is the task of the physicist, whose job it is to say what sorts of bodies are such as to be able to move naturally (such a distinction of course presupposes that the concepts of rest, motion, and space involved are absolute and not relative).

This recalls Aristotle (compare his physical proof of the earth's immobility at *On the Heavens* 2. 14. $296^{a}25$–$^{b}6$), as does the notion that the

astronomers base their demonstrations on incidental properties of the objects in question, such as their sphericity (compare the method of Aristarchus; and cf. *Posterior Analytics* 1. 13. $78^{a}29$–$^{b}12$: we show that the moon is spherical from the fact that it exhibits phases, but its exhibiting phases does not *cause* its sphericity: V.1*a*). Posidonius does not, then, in any real sense espouse an instrumentalist view of astronomy in general. While mathematical astronomy may be principally concerned with the elaboration of predictive models, it is not a matter of indifference (much less of senselessness) whether or not they are true; it is just that in order to determine their truth the mathematician needs the help of the (astro-)physicist (see further XI.1*a*; VIII.5*a*).[26]

(c) Antiochus of Ascalon

Antiochus (*c.*130–67 bc) is a pivotal figure in the development of Middle Platonism. He was for a long time the pupil and associate of Philo of Larissa, the last leader of the sceptical Academy, whose lectures Cicero attended in 81 bc . Philo himself advanced a considerable distance towards modifying the scepticism of Carneades, allowing that certain things could indeed be evident: the dispute between Stoicism and Scepticism was already petering out. However, Antiochus quarrelled violently with Philo, probably around 87 bc ,[27] probably in regard to the question of the unity of the Academic tradition. Philo, apparently, saw it as a seamless tapestry unfolding from Plato through the early Academy to Arcesilaus and Carneades. Antiochus, on the other hand, considered the latter two to mark a profound break with the earlier philosophers (in whose ranks he also numbered Aristotle and Theophrastus: cf. Cicero, *Academica* 1. 15–17); and Antiochus initiated a counter-revolution among the adherents of Platonism, back to what he at least took to be its earlier, pristine orthodoxy.

However, there is more to Antiochus' irredentist syncretism; he leaned heavily towards some aspects of Stoic theory, notably in his adoption of the view that there were two fundamental principles, the active and the passive, with fire and air being active, water and earth passive (ibid. 1. 26; cf. 1. 6; see also VII.1*b*). Equally, matter is without quality, and fit to receive all other forms, as well as being infinitely divisible (ibid. 1. 27). On the other hand, he appears to have borrowed from the Atomists the view that interstitial void is necessary for motion (1. 27). Even his account of providence, as an immanent, intelligent force which permeates and unifies the whole (28–9), is more Stoic than it is Platonic, as is his general leaning

towards materialism (cf. VII.1*b*, 3*a*). Cicero's account of his doctrine of fate is worth quoting in this regard:

363 this [i.e. providence] they sometimes call necessity, because nothing can be otherwise than has been ordained by it as a fated and immutable sequence of eternal order;[28] but they sometimes call it chance, because it brings about many unforeseen and surprising things for us on account of their obscurity and ignorance of the causes. (1. 29)

Antiochus apparently allowed no room either for an independent Demiurge, or for transcendent Forms (he accepts Aristotle's criticism of the Platonic view: 1. 33); if he allowed for forms in any sense at all, he will have made them depend for their existence simply upon their being conceived, in effect assimilating them to the Stoics' common notions.[29] And he was able to treat the Stoic epistemology of infallible perception as being simply a development of the original Academic line (1. 43; cf. 35–42). All in all, Cicero's judgement (2. 132) that, but for some minor changes of doctrine, Antiochus would have been a Stoic, appears justified.

We know most of what we know of Antiochus via the philosophical writings of Cicero, who was, at least some of the time and in some respects,[30] a follower of his (he heard him lecture while visiting Athens in 79 bc). It is not unlikely that his curious *Topics*[31] is derived from a work of Antiochus'; thus at *Topics* 58–66 he reproduces what is probably Antiochus' account of causation:

364(i) there are two classes of causes, one which by its own force invariably produces what is subject to its force, e.g. fire burns,[32] the other which is not essentially such as to produce but without which it cannot be produced, e.g. as bronze might be called the 'cause' of the statue because it cannot be produced without it. (ii) In this class of prerequisites some are quiet, inactive, inert in a way, such as place, time, material, tools, and others of this class;[33] others prepare the way for the production, and add certain things which are in themselves co-operative, even if not necessary, e.g. the encounter provided the cause of love, and love of the crime.[34] From this class of causes in eternal

end p.336

sequence fate is woven by the Stoics. (iii) And just as I have distinguished the classes of causes without which production is impossible, so too can the classes of efficient causes be divided. For there are some causes which clearly produce their effects without assistance, while others require help; e.g. wisdom produces wise men alone and in itself, but whether it also makes them happy alone and in itself is disputed. (Cicero, *Topics* 58–9)

That account is obviously Stoic in inspiration (compare VII.1*c*, **258–9**; *e–f* **263–5**), although the treatment of matter as a prerequisite may be original (cf. VII.1*e*, **263**; XI.3*b–c*, **402–4**). Cicero goes on to underline the fact that confusion among the categories makes for bad argument: just because something is a prerequisite for a certain result does not imply that it contained within itself some genuine causal power (ibid. 60–1: he again invokes the example of the building of the Argo; cf. VII.1*f*). Moreover, some causes are purely physical in their operations, while others operate through the mind by way of beliefs, desires, and dispositions (ibid. 62: this claim need not of course require any modification of psychological materialism); and such causes are not (in themselves at least) inevitable (cf. VII.3*a–b*).

If this is indeed Antiochus, what is striking once more is the extent to which the transcendent, immaterialist aspects of Platonism (again compare **354**) seem simply to have dropped out of the picture. They will not, however, do so for long: and the remaining history of Platonism in antiquity, as it re-emerges in the Middle Platonist synthesis of the succeeding two centuries, before being transformed into the remarkable system of Plotinus, may be seen as turning on the resurrection of the fundamental immaterialism of Platonist thought, as the materialist and mechanist schools which have dominated Hellenistic philosophy go into gradual and irreversible decline. It is this process which takes up the final part of our story.

(d) Seneca on Causes

The last passage showed how Platonism came under Stoic influence, at least in regard to explanation. A century later the Stoic Seneca is more orthodox, insisting that 'the Stoics hold that there is only one cause, that which operates' (*Epistles* 65. 4), endorsing the standard Stoic account of the fundamental Active and Passive principles (65. 2–3; cf.

VII.1*b*). He contrasts this with a standard account of the Aristotelian four causes (65. 5–6), but then remarks, with cheerful anachronism, that 'Plato adds a fifth to these, the paradigm, which he calls the idea' (65. 7). The ideas are contained in the mind of God, taking the form of geometrical figures. Seneca sums up:
end p.337

365 thus there are five causes, as Plato says: (*a*) that out of which, (*b*) that by which, (*c*) that in which, (*d*) that in accordance with which, and (*e*) that for the sake of which. (65. 8)

The attribution of this scheme to Plato is fanciful; but it represents in recognizable form the grafting of a Platonizing concept of non-immanent hypostasized form (*d*) onto the familiar Aristotelian categories. Thus (*a*) picks out the material cause, (*b*) the efficient, (*c*) the formal (although the prepositional formulation used is not well suited to the Aristotelian conception), and (*e*) the final. Platonism, with its commitment to irreducible, ontologically distinct and prior Forms, in accordance with which (hence the prepositional formula) the Demiurge creates the earthly order, clearly requires the addition of an extra category, at least if we are to distinguish between the Form itself and individual instances of the Forms, in the manner suggested by *Phaedo* 102dff. (III.1*b*, **125**).[35]

The use of prepositional formulae to designate the different roles of each category is prefigured in Aristotle (although he is not committed to anything like a formal schematization); and the 'metaphysic of prepositions'[36] becomes the favoured mode of referring to the differing explanatory factors for the remainder of antiquity (see **368** below). Seneca himself rejects this 'crowd of causes' (*Epistles* 65. 11) on the grounds that it is either too broad or not broad enough—too broad if we are concerned with the notion of cause properly so-called, not broad enough if we are to include all the prerequisite factors (if so, time, place, and motion will be causes as well: cf. **364** above).

Seneca's presentation of the 'Platonist' position draws upon the new syncretic Platonism which had developed in Alexandria in the first century bc , and became one of the dominant strands in philosophy in the next 200 years. Properly interpreted, the doctrine maintains, there are no real divergences of belief among the major philosophical schools (of course neither Epicureanism, with its denial of purposive explanation, nor any of the varieties of scepticism could be forced into this syncretizing straitjacket). The report of the 'Platonic' conception of cause in *Epistles* 65 is echoed in the equally syncretic account of *Epistles* 58. 16–22, which ascribes to Plato a sixfold taxonomy of the 'things that are'. The first class is that of generic concepts such as man and animal (58. 16); then comes God (58. 17). The Forms, which Seneca treats as paradigms (cf. **365**(*d*)), are again contrasted with immanent form (18–19). The fifth category is that of ordinary sensible
end p.338

objects, while the sixth contains the quasi-existents, such as space and time, things to which the Stoics refused to ascribe being, although admitting that they were *something* (22).[37]

3. The Emergence of Middle Platonism

(a)Eudorus

Seneca's accounts of 'Platonism' are very likely borrowed from some handbook of Platonic philosophy of the sort which became popular in the succeeding centuries, and of which Alcinous'*Primer on Plato's Doctrines* (§ *e* below) is a prime example. A likely candidate for its authorship is Eudorus, an Alexandrian Platonist who flourished around the middle of the first century bc . It is in what we can reconstruct of Eudorus that we may discern the beginnings of the distinctive Middle Platonist syncretism, as well as what is fundamental to its identity, namely a belief in the transcendent existence of a basic causal principle, the One, from which all else in some way or other derives.

In sharp contrast with the views of Antiochus, Eudorus and his followers insisted upon the transcendence of the Forms, in orthodox Platonic fashion, and, following the examples of Speusippus and Xenocrates (§§1*a–b* above), wove the fundamental Pythagorean categories of the One and the Dyad into their metaphysical scheme of things. Here is Simplicius' report of Eudorus' account of Pythagoreanism:

366in regard to the highest structural principle, it should be said that the Pythagoreans said that the One was the first principle of all, while in regard to the second structural principle, they said that there were two principles of the things which are contrived, the One and the nature contrary to it. And ranged below these are everything which is conceived by way of contrariety, the good under the One, the bad under the nature contrary to it. For this reason these principles are not, according to these men, absolutely primary. For if the one is the first principle of one set of opposites, and the other of the other, then they cannot be common principles of both, as is the One. (Simplicius, *On the 'Physics'* 1. 5. 181. 10–17)

This move, establishing a supreme 'unifying' cause above the basic mathematical principles, is ascribed (by Syrianus, *On the 'Metaphysics'* 166) to the early Pythagoreans Archaenetus, Brotinus, and Philolaus. The ascription is almost certainly false, relying no doubt on the testimony of one of the numerous Pythagorean forgeries that circulated in Hellenistic and later times; but the forgeries themselves are of independent interest, since they

contain the new 'Pythagorean' doctrines. Eudorus must have had access to these, or similar, documents,[38] and approved of their fundamental contentions; indeed, the Pythagorean forgeries may themselves have been created by Alexandrian Platonists of the period anxious to confect a spurious 'tradition' justifying their own doctrinal innovations.[39] Eudorus in fact makes his supreme principle actually responsible for the generation of matter: Alexander, *On the 'Metaphysics'* (commenting on *Metaphysics* 1. 6. 988^{a}10–11, where Aristotle ascribes to Plato two principles, matter and essential form), remarks that Eudorus 'emended' this text to make the One responsible for the creation of matter as well as the Forms. And, as Dillon (1977: 127) notes, there is no independent evidence in any Pythagorean source of the period or earlier for the notion that above the basic division of Monad and Dyad there is a supreme One, a fact which makes it more reasonable to suppose that what Eudorus attributes in **366** to unnamed Pythagoreans is in fact his own view (or at the very least his own innovative interpretation of the Pythagorean tradition).[40]

Eudorus accepted, at least in part, Xenocrates' account of the soul as being generated by the original numbers (Plutarch, *On the Generation of the Soul in 'Timaeus'* 3. 1013b), but he also wished to accommodate the view of Xenocrates' pupil Crantor that the soul consisted of a mixture of intelligible substance and that which is suited to take on impressions of perceptible objects, a dual nature required by its intermediate position between the sensible and the intelligible realms. Eudorus also evidently adopted the view (which was to become standard in later Platonism) that the *Timaeus* was not meant to describe a literal creation of the universe in time, but rather to exhibit the relations of metaphysical dependence that held among its (eternal) parts, a view also endorsed by Eudorus' Alexandrian near-contemporary Philo (*On the Creation of the World* 7–10; cf. Plotinus: XII.2*d*).

(b) Philo of Alexandria

Philo (*c*.30 bc –*c*.ad 45), a Jewish thinker and allegorizing exegete of the Old Testament much of whose voluminous *œuvre* has survived, was not a

Platonist as such; but he was deeply influenced by Greek philosophy in general, and Pythagoreanism and Platonism in particular. Some of his views on the structure and relations of the parts of the universe bear brief examination here. His God is basically the God of the Old Testament: but he is also a transcendent, quasi-mathematical object in the Pythagorean tradition (cf. *On the Creation of the World* 100). He is, in fact, beyond the reach of reason (*Embassy to Gaius* 6), and much of Philo's theology is negative, and as such prefigures the constant struggle of Plotinus to speak intelligibly of the ineffable (XII.2*b*). However, God's existence can be inferred from his effects (Philo subscribes to the argument from design: *On The Special Laws* 1. 33–5; cf. VII.3*a*, **282–3**; XI.3*b*); and while it is hard to say anything directly about his essence, it is possible with intellectual effort to attain at least some understanding of it (ibid. 1. 26–41). The goal is to attain knowledge of 'him who is really real [*ontōs on*: the phrase is Platonic: XII.1*b*, **438**; cf. *Republic* 5. 477a], who is the first and most complete good' (*On the Decalogue* 81: such locutions for God are ubiquitous in his works), something which can, however, be achieved fully only by the most accomplished (*On Flight* 97–100).

One passage from *On the Creation of the World* will suffice to give the flavour of Philo's distinctive mix of Jewish theology and late Hellenistic syncretism:

367 some people [i.e. Peripatetics], wondering at the world rather than the world's Maker, say that it is ungenerated and eternal, while falsely and ignorantly ascribing complete inactivity to God. . . . Moses, having attained the pinnacle of philosophy, and being instructed in the greatest and most essential aspects of Nature, knew that most necessary among things that are were an active cause and a passive, the active being the purest and most pristine Mind, greater than virtue, knowledge, and the Good and Beautiful themselves, while the passive is without life and of itself motionless, but when set in motion and shaped and animated by Mind, changes into the most perfect of works, this world. (Philo, *On the Creation of the World* 7–9)

We may discern in this elements of both Platonism (from *Timaeus* and *Philebus* 30a–e: cf. III.4*a*) and Stoicism, as well as a bow in the direction of Jewish orthodoxy: the world is indeed created (although not in time), for otherwise it would not exhibit divine providence (*On the Creation of the World* 10). Whether or not the world is created absolutely *ex nihilo*, matter being created as well (*à la* Eudorus), or whether matter in some sense preexists the creation (in orthodox Platonic fashion), is unclear; but since the creation is not creation *in* time (since time too is created, again in line with the *Timaeus*: 38b), the question may be of little import: at all events, Philo

end p.341

affirms that there can have been no time when matter was not in its divinely organized state (*On Providence* 1. 8).[41]

On the Creation of the World 16–19 envisages a double divine creation, with God first producing the intelligible cosmos and then employing it as a paradigm to create the sensible world, an adaptation of the *Timaeus* picture (III.4*c*; cf. **371**). The active principle of **367** operates in the world by way of *logos*, the divine reason permeating all things in the manner of the Stoics' *pneuma*, which is described as the location of the Forms (which for Philo too are numbers: *On the Creation* 102; *Who is Heir to Divine Things?* 156). The Forms become, through the activity of the *logos*, *spermatikoi logoi*, the Stoics' seminal principles (*Embassy* 55; cf. *Allegories of the Law* 3. 150; *Who is Heir?* 119); these operate to form and maintain the furniture of the perceptible world, dividing up and imposing order upon the unarticulated basic matter (ibid. 133–236). This *logos* is also conceptualized as God's instrument, that through which the world is created (*Allegories of the Law* 3. 96). God is not identical with this instrumental *logos*; rather he is distinct from it, being its Cause. A passage from *On the Cherubim* makes this clear, as well as exemplifying Philo's acceptance of the 'metaphysic of prepositions' (n. **36** above):

368 God is the cause (*aition*) not the instrument; and what is generated is generated not through an instrument but by the cause. Many things must coincide for the generation, the by which, the out of which, the through which, and the for which; and the by which is the cause (*aition*), the out of which the matter, the through which the instrument, and the for which the explanation (*aitia*). (Philo, *On the Cherubim* 125)

The insistence that the efficient cause is the only true cause is of course Stoic (cf. §2*a*, **364**; §2*b*; VII.1*a*; XI.3*b*), the willingness to allow the other categories explanatory elbow-room now simply being par for the syncretist course (note, however, that Philo's particular distinction between *aition*

end p.342

and *aitia* is quite different from that ascribed to Chrysippus: VII.1*c*, **255**). None the less, the emergence of the instrument as a distinct causal category is worth remarking as another example of the proliferation of causal categories abominated by Seneca; and Galen was to take up and develop the notion (XI.3*a–b*, **401–2**; cf. XII.3*b*, **476**, 3*d*, **487–9**). Philo continues with a stock account of the causal categories involving the standard example of a house built by a builder, from materials, through the employment of tools, for the sake of protection (ibid. 126; cf. Aristotle, *Posterior Analytics* 2. 11. 94^{b}9–10). However, this is merely an illustration:

369 moving on from particular constructions examine the greatest house and city, this world. You will find that its cause is God by whom it was generated, its matter the four elements out of which it was composed, the instrument the *logos* of God through which it was constructed, and the explanation of the construction the Creator's goodness. (Philo, *On the Cherubim* 127)

The case of 'explanation' (*aitia*) seems not to be quite on all fours with either the general prepositional formula or the construction example. God's goodness does not seem to be a final cause of the universe's creation, and Philo is not always as careful as he might be in his deployment of such categories, as his treatment of the relations between God, the *logos*, and the other powers shows. A passage from *On the Creation* may ease the tension between **368** and **369**:

370 the power that created the world had for its source true goodness. For if anyone should want to seek the explanation (*aitia*) for the sake of which this whole was created, it seems to me he would not miss the mark in saying . . . that the Maker and Father was good; for the sake of this he begrudged nothing of his own excellence of nature to that substance [i.e. matter] which has in itself nothing fine but is capable of becoming all things. For it was in itself disorderly, without quality, inanimate, irregular, full of dissimilarity, disharmony, and discord. But it could turn about and change in the direction of the opposite, best things, namely order, quality, life, regularity, identity, the harmonious and concordant, everything of the better form. (Philo, *On the Creation of the World* 21–2)

Here Philo perhaps suggests that God's goodness is itself the reason why the world has the final-cause explanation that it has, namely that it was constructed for the good. Supreme among the divine powers associated with (and derived from) the *logos* are Goodness (*agathotēs*), the creative principle, and Sovereignty (*exousia*), the ruling principle (*On the Cherubim* 27–8). These Powers Philo describes as 'shadows' of God, although not strictly equivalent to the transcendent supreme being (*On Abraham* 119–24); thus God may sometimes be appreciated as a kind of trinity, but to the initiated he can be seen independently of the numbers (i.e. the Forms), and the two principles

(assimilated to the Dyad: ibid. 124) in a pure and unmixed Monadic condition: this too foreshadows key elements in the Plotinian scheme and its associated spiritual exercises (XII.2*c*). A complication is introduced by Philo's adoption of a fundamental female principle of creation (cf. Xenocrates: §1*b*, **352** above) which he labels Wisdom (*sophia*), and is sometimes assimilated to the Greek notion of Justice (*dikē*). *Sophia* is sometimes equated with *logos*, but sometimes appears to precede it: *sophia* is that 'through which the universe was generated', i.e. she functions as an instrumental cause (*On Flight* 109; cf. *The Worse Attacks the Better* 115–16; *On Drunkenness* 30–1).

In allegorizing Genesis 18: 6, where Abraham tells Sarah to mix three measures of wheat flour and bake cakes, Philo interprets the measures in cosmological terms:

> **371** one measure is that by which the incorporeal and intelligible world was constituted. The second measure is that by which the perceptible heaven was established in the fifth element, attaining to a more wonderful and divine essence, unaltered and unchanging by comparison with the things below and remaining the same. The third measure is the way in which sublunary things were made out of the four powers, earth, water, air, and fire, admitting generation and corruption. Now the measure of the incorporeal forms by which the intelligible world was constituted must be said to be the oldest of causes; and the cause of the fifth, perceptible and circular essence, which the heaven has had allotted to it, is the creative power of the Existent One, for it has found an imperishable, pure, and unalloyed blessing in obtaining an immortal and incorruptible portion. But the sovereign power is the cause of sublunary things, those that are subject to change and alteration because they participate in generation and corruption. . . . So that truly and strictly speaking, God alone is the measure of all things,[42] both intelligible and sensible, and he in his Oneness is likened to a triad because of the weakness of the beholders. (Philo, *Questions on Genesis* 4. 8, trans. after Marcus[43])

It is worth noting Philo's admission of the Aristotelian fifth element (again compare Xenocrates, Frs. 264–6 IP: §1*b* above) into his fundamentally Platonic scheme, a scheme in which God properly so-called is at a remove from either of his executive powers, which themselves stand in a hierarchy, the creative being superior to the sovereign (cf. *On Flight* 97–100); they both, however, owe their existence ultimately and pre-eminently to God as the first cause, as does everything which in turn depends on them (*On the Confusion of Tongues* 123–4).

It remains to say something about Philo's conception of freedom. Given the last remarks, it is plain that in some sense the entire world, including its human denizens, is causally dependent on God. What is more, the universe is providentially ordered (Philo wrote a treatise *On Providence*, the bulk of which survives only in Armenian, devoted to proving this; but the doctrine is ubiquitous: cf. e.g. *On the Creation* 171–2). Yet Philo also stoutly maintains that human beings are free.[44] In the course of describing the increasing excellence of living things as one progresses up the *scala naturae*, having established that intelligence is humankind's peculiar gift he writes:

372 this part of the soul was not fashioned from the same elements as those which formed the others, but was allotted a purer and better substance, out of which the divine natures[45] were made; for this reason it seems reasonable that mind is our only immortal part. For the Father who generated it considered it alone fit for freedom, and slackening the bonds of necessity allowed it free rein, endowing it with that portion of the will which is his most appropriate and proper possession which it was able to receive. For he handed over the other animals, in whose souls that which is set aside for freedom—the mind—does not exist, yoked and bridled to humanity as servants to masters. But men, allotted a self-willed and self-determined intelligence, whose actions are for the most part made as a result of choice, are reasonably blamed for injustice committed intentionally and praised for the voluntary good deeds. The successes and failures of plants and other animals are neither praised nor blamed, since their movements and changes in either direction occur without choice or volition. Only the human soul has received from God voluntary movement, and is in this respect made most like him; and having been set free, as far as possible, from that hard and ruthless mistress Necessity, justly meets with accusation because it does not honour that which set it free. (Philo, *On God's Immutability* 46–8)

Philo, then, asserts that the human will is free as a result of the faculty of choice. But he does not argue for this assertion; nor does he anywhere show how it can be made compatible with his doctrines of the causal priority of God, of divine providence, and of a pervasive *logos*, to which elsewhere in the same treatise he attributes the human capacity for right and wrong action (ibid. 128–39), or with the view of *On the Cherubim* 128 that we are merely the instruments through which God's providence is worked out (cf. **368–8**). This was to become the standard problem for later Platonist theorizing on the issue (§ *c*; cf. XII.2*d*, **459–64**), not to mention later Stoicism as well (cf. Epictetus, *Discourses* 1. 1, 14, 16, 3. 17; cf. VII.3*a*).

end p.345

Plutarch and Pseudo-Plutarch

Allowing for disagreements and obscurities, we can now discern the general outlines of the Middle Platonic synthesis. Plutarch of Chaeroneia (*c*.ad 45–125), a voluminous author much of whose work survives, is (not- withstanding occasional Carneadean leanings) a typical representative of the emerging philosophical consensus. His fundamental principles are the **Pythagorean** pair of the One and the Indefinite Dyad, assimilated to an all- knowing, providential, immutable God, and formless matter respectively (*On the E at Delphi* 392e–393d; *On Isis and Osiris* 351d, 382b; *On the Obsolescence of Oracles* 423d, 428f). God intervenes in the world via the *logos* (*On Isis* 373a–b); and the Forms exist in his mind (*On Divine Vengeance* 550d). The formless matter, assimilated to the Egyptian deity Isis, and the female principle in general (compare Philo's *sophia*: § *c* above) is also an irrational World-Soul (*On Isis* 351e–f, 372e; *On the Creation of the Soul* 1026e–f), which is responsible for evil, since such a thing cannot be attributed to the beneficence of God (*On Isis* 369a–e):

373 for if (*a*) nothing is such as to come to be uncaused, and (*b*) the good would not provide a cause for evil, nature must have a particular generation and origin of evil, just like that of the good. (Plutarch, *On Isis* 369d)

Premiss (*a*) of this argument is our old friend CP1*a* (I.3*d*; cf. P1–2: III.1*b*), while (*b*) is evidently related to the Principle of Causal Synonymy (PCS: I.3*c*–*d*), and is a special case of Plato's P3 (III.1*b*), that causes cannot bear properties incompatible with the properties they induce. This Manichaean dualism was taken up by neo-Pythagoreans, such as Numenius (Frs. 11, 50, 52 dP:[46] cf. XII.1*b*, **434–6**), but was definitively rejected by Plotinus (XII.2*a*–*b*).

Plutarch also evinces a keen interest in matters cosmological. His curious text *On the Face in the Moon*, which is part astronomical disputation and part mythical cosmic demonology, takes the form of a debate between representatives of various schools concerning the nature and structure of the universe, and most particularly (unsurprisingly) the moon. The dramatic representative of Plutarch's Academic views (his brother Lamprias) is concerned, among other things, to support the view that the moon is fundamentally earthy in physical constitution. In order to do this, he must deflect some obvious objections from Stoics and Peripatetics alike:

374 helping the moon not to fall is its motion and the speed of its revolution, just as things placed in slings are prevented from falling by the circular orbit.[47] For natural motion controls each thing, unless it is diverted by something else. For

end p.346

this reason, the moon's weight does not control it, since its force is disrupted by the orbital motion. In fact it might make more sense to wonder at it if it were completely stationary and unmoved like the earth. As things are, while the moon has the best reason for not moving towards us, since the earth has no share of any other motion, it is reasonable to suppose that it is not moved by weight alone. (Plutarch, *On the Face in the Moon* 923c–d)

That remarkable passage shows just how far people were already prepared to go in rejecting the Aristotelian paradigm; and the concept of the moon's natural centrifugal tendency in rotation balancing its natural inclination as a heavy object to fall prefigures the scientific revolution of the seventeenth century (cf. Plotinus' view: XII.2*d*). Lamprias goes on to point out that since, as Aristarchus has shown (§1*b* above), the moon is very close to the earth compared with the other heavenly bodies, it is not unreasonable that it should resemble it in structure (ibid. 925a–f).

Moreover, for the Stoics at least (who suppose the finite cosmos to be surrounded by an infinite void), there is no centre to the universe as a whole, and hence no reason to think that it, rather than the earth itself, possesses the attractive power (925f–926b; cf. V.2*b*):

375 the totality, being infinite, having neither beginning nor end can have no middle either. . . . He who says that the earth is in the middle not of the totality but of the cosmos is naïve if he does not think that exactly the same difficulties apply in the case of the cosmos as well; for the totality has left no middle for the cosmos, but having neither hearth nor home it moves in an infinite void to no proper place; or, if it has come to a stop by finding some other reason for resting and not because of the nature of its place, similar conjectures may be made regarding the moon as well as the earth, namely that the one remains unmoved here and the other moves there as a result of their different souls and natures rather than a difference in place. (925f–926a)

Thus the different motions of the heavenly bodies are ultimately to be ascribed to their souls, although their particular trajectories may also involve, as **374** suggests in the case of the moon, the influence of their particular material structure. Further empirical nails are driven into the coffin of the doctrine of natural place by a consideration of the fire to be found within the earth in volcanoes, and other such natural phenomena (926c–927a). It is then argued that if everything in the cosmos were organized purely by the natural tendencies of things to seek their natural places, there would be no room for providence in the world (927a–d). Finally, the idea that there is a single natural motion for each thing is rejected: rather what is natural in the proper sense is motion in accordance with reason and soul (927d–928d).

In his rebuttal of the Stoic account of the structure and illumination of the moon (cf. Posidonius, Frs. 125–6 EK), Plutarch also shows himself well

acquainted with contemporary developments in optics (cf. V.3*c*; XI.1*a*). The Stoics resist the view that the moon shines by reflected light on the grounds that, being spherical, and given the fact that the angle of reflected light equals that of its incidence, it should then appear only as a point (*On the Face in the Moon* 929e–930a; cf. Cleomedes, *On the Heavenly Bodies* 2. 4; they also hold that, since it is tenuous in substance, it should not reflect at all). They take its illumination rather to be the result of a mingling of the sun's rays with the particular lunar substance (ibid. 2. 4). Plutarch accepts that the angles of incidence and reflection are equal (although perhaps only for the sake of argument: he describes it as ‘neither self-evident nor an admitted fact’, supposing that it does not hold in the case of curved mirrors: *On the Face in the Moon* 930a–d; cf. §4*a* below), but points out that such perfect reflection should only occur from a perfectly smooth body, which fact itself supplies a reason for supposing the moon to resemble the earth in being earthy and uneven (ibid. 930d; various further objections are then raised to the Stoic account: 930e–931c).

The behaviour of light when it strikes the earth is similar to its behaviour when it strikes the moon, but utterly dissimilar to what it does in air; hence, since like effects should be produced by like causes, we must conclude that the moon is solid and earthy in structure (931b–c). In the same vein, and employing the same causal principle (‘if the effect is similar the causes are similar’: 932a; cf. P1 and P2: III. *b*), Plutarch argues to the same conclusion from the fact that the effect of a lunar eclipse (in which the earth deprives the moon of light by interposition) is identical in type to that of a solar eclipse (in which the moon similarly deprives the earth of light); hence their substances must be (at least in so far as their reaction with light is concerned) the same.

The heavenly bodies, then, are rational, and their motions are guided like everything else in the world by providence. None the less, for Plutarch as for all Platonists the human will is free. As far as the relation between freedom and divine providence is concerned, his position represents the emerging orthodoxy (cf. Philo, **372**, § *b* above): Plato, he claims,

376 is always touching on the three causes, since he was first (or at least most consistent) to observe how fate is naturally inclined to mix and intertwine with chance, while what is up to us (*to eph'hēmin*) in turn combines with one or other, or both at the same time. So here he has admirably clarified what power each has in our affairs, assigning the choice of lives to us . . . while associating with the compulsion of fate the good life of those choosing well and the opposite state of those who choose badly. And the fall of the randomly scattered lots introduces chance which determines many things in our lives, along with the upbringing and society in which each finds himself. (Plutarch, *Table Talk* 9. 5. 740c–d)

end p.348

Chance, although a cause (as Aristotle held: *Physics* 2. 4. $195^{b}30$; cf. IV.2*b*, §4*b* below), is not itself caused; fate is simply the natural order of things, how things will turn out if certain antecedent events take place: but those antecedent events (or at least those which fall under the purview of our choice) are not themselves subject to fate or determination. Furthermore, God's providence is distinct from, and may on occasion override, fate by interfering in the natural order of things (this interference may be permanent, as it is in the case of some of the heavenly bodies, notably the earthy moon, which is held aloft against its nature by 'reason', because it is better thus: *On the Face in the Moon* 927a–d; this of course amounts to a rejection of the principles of Aristotelian physics, that eternal motions must be given a natural explanation: V.2*g*, **198–9**; cf. **192**).

This concept of fate as a sequence of hypothetical consequences is elaborated in the text *On Fate* wrongly attributed to Plutarch. But although the attribution is certainly false, the text presents an orthodox Middle Platonic account, and it probably dates from the first half of the second century ad .[48] The author begins by distinguishing the substance of fate (which is identified as the World-Soul: cf. III.4*a*; § *d* below) from its activity; from Plato he derives three characterizations of the latter, respectively 'a divine *logos* untransgressible because of unavoidable causation'; 'a law in conformity with the nature of things, in accordance with which events unfold', and 'a divine law in accordance with which past, present, and future events are intertwined' (*On Fate* 568d; these definitions are derived, somewhat fancifully, from *Phaedrus* 248c, *Timaeus* 41e, and *Republic* 10. 617d respectively). Only the first seems to involve anything like determinism; but our author backs away from such an implication. In the course of expressing a belief in an eternal, cyclical recurrence of events (a belief shared by both Pythagoreans and Stoics), he writes:

377 the arrangements [sc. of the heavenly bodies] that follow in sequence and depend upon one another will follow in a dependent way, each bringing along its own events (however, note for clarity's sake . . . that it is not as a result of the heavenly bodies as universal causes of things that I am now writing this and you are doing what you happen to be doing). And so when the same cause returns, we shall, becoming the same people, do the same things, as will everyone else. (pseudo-Plutarch, *On Fate* 569b–c)

Thus the heavens are not to be held responsible for everything—on the other hand everything will turn out the same in the succeeding cycles (presumably those of the Great Year: VIII.5*b*, **316**). It is not clear what to make of this. At first sight it resembles the Chrysippean account

(VII.2*a*–*b*); yet this is precisely what our author is out to rebut. Fate operates like a political law; it is hypothetical in structure, decreeing penalties that will ineluctably follow if we transgress (ibid. 569d–570a); this, too, is the form of fate, and, incidentally, the meaning of the first definition derived from the *Phaedrus* (570a–b). Our author then tries to make clear the relations between fate thus conceptualized and providence on the one hand and chance on the other, holding that ‘everything conforms to fate’ is true only in the sense that nothing happens contrary to it, not in the sense that fate is itself the actual determinant of all events (570b–e; cf. Ptolemy's attitude to astrology: XI.1*b*, **395–6**).

It remains, then, to determine the scope of what is up to us, chance, the contingent, and the possible, all things which fall outside the direct determination of fate (570e–f), a project which occupies 570f–572f, and which is very largely Aristotelian in inspiration. There follows (572f–574d) an account of providence and its relations to fate (it is prior to it, at least in its highest manifestation: 574b). The upshot of all of this is that although we are in control of our choices, those choices themselves have inevitable results; but our author never specifies quite how this is supposed to be the case. He allows that the doctrine of providence might appear to entail determinism (574c), but rather lamely concludes that he does not wish to pursue the issue any further, preferring to think of the antecedent causes as existing alongside fate (574c–d). For all its shortcomings, it is clear from this treatise that Platonists felt themselves forced to engage with the Stoic arguments regarding fate and determinism. It is, I think, also clear that their responses to it owe more to wishful thinking than they do to cogent argument, as is so often the case with questions of the will.

(d) The Handbook of Alcinous

There survives complete from the second century ad [49] a handbook of Middle Platonism attributed in the manuscripts to one Alcinous. It was argued in the last century by Freudenthal that this was a corruption for Albinus, a well-known Platonist of the middle of the second century, but modern scholarship now tends towards the view that the

manuscript ascription is correct.[50] Nothing, however, is known of this Alcinous. His *Primer on Plato's Doctrines* is a work of little, if any, originality; but it usefully preserves a compendium of mature Middle Platonism, and as such merits brief notice here.
end p.350

After a short account of logic and epistemology, in which he contrives to father on Plato much that is really owed to Aristotle, Alcinous turns in chapter 8 to the fundamental principles of the universe. He treats of matter first, relying heavily on *Timaeus* 49a–52d, the only obviously Aristotelian elements being the word for matter itself (*hulē*) and its characterization as being neither bodily nor incorporeal, but potentially body (cf. Aristotle, *On Generation and Corruption* 2. 1. 329[a]33). Chapter 9 deals with

378 the other principles, the paradigmatic [i.e. that of the Forms] and that of the Father and Cause of all the gods. Considered in relation to God the idea is his thought, to us the first intelligible, to the sensible world the paradigm, to itself substance. (Alcinous, *Primer on Plato's Doctrines* 9. 163. 12–17)[51]

Forms are defined, following Xenocrates (**354**), as the paradigms of things which exist naturally, and these exclude artefacts, unnatural states such as disease, individuals, or relational properties such as larger and smaller; Forms are 'the eternal and perfect ideas of God'. Alcinous next proves (to his own satisfaction at least) that Forms must exist on four grounds. God, being an eternal intellect, must have eternal ideas; matter, being essentially without measure, needs to have form imposed upon it; if knowledge is distinct from opinion, so too will be its objects; but most interesting from our point of view is the claim that

379 if the world is not as it is by chance,[52] but has been generated not only out of but also by something and additionally in accordance with something; but in accordance with what else if not a Form? Consequently, there will be Forms. (ibid. 9. 163. 38–164. 1)

That argument is, as it stands, enthymematic and unconvincing. But in addition to confirming the grip of the 'metaphysic of prepositions' (n. **36** above), it exhibits in compendious form the fundamental considerations the Platonists of the period brought to bear against the views of their opponents. That the universe could be merely a matter of chance as the Atomists would have it, or that it could have a purely material explanation, is dismissed out of hand. But then, if further explanatory categories are required, we must import both an agent, and something the agent had in mind: and the latter could only be Forms. That of course will convince no urbane Aristotelian (or indeed pure mechanist); but some of the considerations involved, notably in favour of postulating a conscious creation at the hands of a divine Craftsman, will be taken up (in a greatly more sophisticated form) by Galen and others (XI.4*a*).
Chapter 10 turns to God, intriguingly distinguishing mind from soul:
end p.351

380 since mind is better than soul, and that which thinks everything simultaneously and eternally in actuality is better than mind in potentiality, finer even than this will be its cause and whatever might exist even higher than them; and this would be the primary God, which is the cause of eternal activity for the mind of the entire heaven. And it [i.e the primary God], being itself motionless, causes activity in accordance with the latter [i.e. mind], as the sun does in accordance with sight when anyone looks upon it, and as the object of desire causes desire being itself motionless: thus will this mind move the mind of the entire heaven. (10. 164. 18–27)

This contains Aristotelian echoes (notably in the ultimate object's being an object of desire: Aristotle, *Metaphysics* 12. 7. 1072^{a}23–^{b}4, **204**; V.2*g*); and Alcinous goes on to describe the primary mind as eternally contemplating itself. The World-Soul is of course thoroughly Platonic (III.2*a*, 4*a*), but the World-Mind (at least as a separate entity) seems proprietary to Alcinous (or more probably to his source).

God is eternal, ineffable, self-complete, eternally complete, and entirely complete; He is also divinity, substantiality, truth, symmetry, and good. He is called 'father'

381 in that he is the cause of everything and sets in order the heavenly Mind and the World-Soul in accordance with himself and his own thoughts. For he has filled everything with himself in accordance with his own will, awakening the World-Soul and turning it towards himself, being the cause of its Mind; and when this has been set in order by the Father, it sets in order everything natural in the world. (10. 164. 40–165. 4)

Thus the Mind is part of the World-Soul, rather than being a completely distinct entity; and the Primary God is also an intellectual being in Aristotelian fashion. But Alcinous' model is clearly a precursor to Plotinus' three hypostases (XII.2*a*).

Finally among the principles Alcinous includes the qualities. These are incorporeal, for (among other reasons) if they were bodily, then two or more bodies would occupy the same place, which is absurd. If qualities are incorporeal, then what creates them must be incorporeal as well; and what is active must be incorporeal too, since it contrasts with the passivity of body. All of this is in sharp contrast to the materialism of the Stoics (VII.1*b*), for whom qualities were bodies disposed in a certain way, and for whom all causation was material (VII.1*b*, *c*, **256–7**; the Stoics did, of course, treat effects as incorporeal *lekta*); but it is entirely representative of the developing tenor of the times. From now on, with a few distinguished exceptions, notably Alexander (§4) and Galen (see XI. 2–5) causal agency is generally conceived of as fundamentally a mental phenomenon, distributed by mysterious, insubstantial means throughout the entire cosmos, and of which the apparent causal interaction of bodies is merely a pale and misleading shadow (see XII.2*b*, 3*a*).

4. The Revival of Aristotelianism: Alexander of Aphrodisias

(a) Background

The history of the Lyceum after Strato's death (V.3*c*) is obscure (although the Peripatetic research programme was carried on by scientists such as Aristarchus, Strato's pupil: cf. X.2*b*). In the second century bc , one Critolaus tried to revive Aristotelianism; he is reported by Philo (*On the Eternity of the World* 70: §3*b* n. **41**) to have argued for the world's indestructibility (against the Stoics) on the grounds that whatever causes itself to exist does so necessarily, and hence eternally, while the world causes itself to exist, since it causes everything else to exist (cf. Plotinus, *Enneads* 6. 8. 13–14).

The revival in the fortunes of Aristotelianism began with the canonical edition of Aristotle's esoteric treatises made in the middle of the first century bc by Andronicus of Rhodes. But Andronicus was more than merely a skilled philologist; he also inaugurated the tradition of commentary, often of an innovative revisionary kind, on Aristotle's texts which was to become the main medium of philosophical activity in later antiquity. Most significant from our point of view is his apparent insistence that, although all motions are in some sense caused by external movers (cf. Aristotle, *Physics* 8. 4. $255^{b}134$–$256^{a}3$; V.2*b*–*c*), strictly speaking the real source of the motion or change is to be found within the changing thing. We know from Simplicius (*On the 'Physics'* 3. 3. 440. 12–17) that he adopted a particular reading of Aristotle's *Physics* 3. 3. $202^{a}14$ in order to bolster this contention;[53] ten pages later (*On the 'Physics'* 3. 3. 450. 16–26), Simplicius quotes with approval Andronicus' doctrine that 'the nature [sc. of each thing], having itself been predisposed, disposes that thing from the inside in each type of change'.

Shortly thereafter the Peripatetic Xenarchus argued, on fundamentally Aristotelian grounds, against the existence of the fifth element, holding that it was not in fact contrary to nature for fire to move in a circle. The natural motions of the sublunary elements are rectilinear only when they are not in their natural places; in their natural places they are naturally either at rest or move in a circle (Simplicius, *On 'On the Heavens'* 1. 2. 13. 21–8).[54] At roughly the same time Nicholas of Damascus (64 bc –*c*.ad 20)

produced, among a vast and polymathic *œuvre* that included comedies, tragedies, and a *Universal History* in 144 books, volumes entitled *On the Gods* and *On the Totality*, and a treatise entitled *On the Philosophy of Aristotle*, fragments of which survive, mostly in Syriac.[55]

The latter is an intelligent digest of Aristotle's views, culled from the *Physics*, *Metaphysics*, *On Generation and Corruption*, and *On the Heavens*. Two points are worthy of note. First, in detailing the various senses of *archē* after the manner of *Metaphysics* 5. 1, Nicholas writes: 'one is the origin of the fact that something is, e.g. essence, matter, nature, and semen' (*On the Philosophy of Aristotle* 2. 3); as Drossaart Lulofs notes (1965: 100–1), Nicholas must be thinking of semen as the efficient cause, with nature as the final cause. Second, in his epitome of the doctrine of the four causes (*On the Philosophy of Aristotle* 10. 1–6), Nicholas subscribes to the 'metaphysic of prepositions' (n. **36** above): 'these causes are four: that out of which, that in virtue of which, that from which, and that for the sake of which' (10. 6), although again his practice, at least in regard to the third category, that of efficient causation, is slightly unorthodox (compare §2*d*, **365**; §3*b*, **368–9**; §3*d*, **379**).[56]

Also worth mentioning from the period is the short text *On the World* falsely attributed to Aristotle.[57] It takes the form of a letter addressed to Alexander the Great, eulogizing the beauty, nobility, and wondrousness of the cosmos. The treatise is Peripatetic in general doctrine, embracing among other Aristotelian meteorological doctrine the theory of exhalations (*On the World* 4. 394^{a}9–22; cf. Aristotle, *Meteorology* 1. 4–12; V.2*d*), although its god is no mere disengaged Aristotelian Prime Mover: rather he is the constant preserving force (the author employs the Stoic term *aitia sunektikē: On the World* 6. 397^{b}9; cf. VII.1*b–c*) of the cosmos, and the creator of everything within it, albeit indirectly and by remote control (6. 397^{b}20–398^{a}6). This god is not Aristotle's, but neither is it the immanent, all-pervasive, material god of the Stoics.

Perhaps a century later[58] Hero of Alexandria, mathematician and inventor,[59] shows some signs of Peripatetic influence. The introduction to his *Pneumatics* recalls, albeit less unmistakably than it is sometimes claimed, Strato's physical theories. His *Catoptrics* (which survives only in Latin)

end p.354

discusses, among other things, why light scatters when reflected from apparently smooth objects, concluding that it is only in highly polished surfaces that the full mirror-effect is possible, since otherwise the light is dispersed by minute pits in the surface; and he is aware of the fact that the angle of reflection equals the angle of incidence (something in fact known a good deal earlier: compare Plutarch, *On the Face in the Moon* 929e–930e: §3*d*). Intriguingly, he attributes this to its being *reasonable* that things should be thus; Olympiodorus (*On the 'Meteorology'* 212–13) glosses this as being the result of nature's doing nothing in vain, as it would if the light were to follow any other path (since any other path from object to eye would be longer). Hero thus anticipates the modern physical principle of least action, but gives it a distinctive Aristotelian metaphysical foundation (Damianus, commenting on the same passage, directly ascribes to Hero the distinct, albeit compatible and equally Aristotelian, view that nature behaves thus 'because she is not about to lead our sight in vain': *Optics* 14; cf. X.2*b*; XI.1*a*).

Virtually nothing else of substance survives of the Aristotelianism of the period until the time of Alexander of Aphrodisias. Alexander was the leading Peripatetic of his day, and the author of numerous commentaries on the works of Aristotle, many of which survive. Little is known of his life. He was appointed to the chair of Peripatetic philosophy in Alexandria, which had been established in ad 176 by Marcus Aurelius, probably around the turn of the century: our sole evidence for his dates derives from the fact that he dedicated his *On Fate* to the emperors Severus and Caracalla, whose joint principate lasted from 198 to 209. Chronology makes it unlikely, although not impossible, that he is identical with the Alexander of Damascus whom Galen mentions as the official teacher of Aristotelian philosophy in Athens (*On Anatomical Procedures* 2. 218), and with whom he quarrelled (*On Prognosis* 14. 627–9).[60]

Alexander's intellectual masters included Herminus (against whose championship of the Aristotelian Prime Mover Galen wrote a treatise which Alexander later refuted[61]), and Sosigenes, a Peripatetic interested in physical and astronomical problems.[62] He was not the first to write

end p.355

Aristotelian exegeses: we know of commentaries written by Adrastus and Aspasius (fl. *c*.ad 100; they were discussed in Plotinus' seminar: XII.1*a*, **433**), one of which (Aspasius on the *Nicomachean Ethics*) partially survives. But he was certainly one of the originators of the great exegetical tradition that came to dominate late antique and medieval philosophy; and his works on the *Physics* and *On the Heavens*, even though they do not survive, exercised a great influence on his successors (Simplicius in particular quotes from them at length).

In addition to the commentaries, there are some surviving monographs on physics, ethics, the soul, and fate, some of doubtful authenticity; and he wrote an *On Providence*, which survives at least partially in Arabic, and defends the view that providence consists in the beneficial effects the orderly motions of the heavens have on the sublunary world (see also *Questions* 2. 3, 21).[63] These texts are frequently polemical in tone, and their principal targets are the Stoics, and as such we have already met *On Mixture* (VII.1*b*) and *On Fate* (VII.1*d*, 2*a*, *d–e*); but it seems plausible that his monographs were written to fill the gaps in those areas which, in Alexander's view at least, had not been sufficiently dealt with by Aristotle. Thus both *On Fate* and *On Providence* essayed Aristotelian treatments of themes which had only become central to philosophy with the development of Stoicism. In addition to these texts there survives Alexander's extensive *On the Soul*, a paraphrastic interpretation of Aristotle's treatise, and the heterogeneous collection of short essays on subjects ranging from psychology through physics to fate and providence, which Bruns (1887) lumped together under the compendious title of the *Mantissa*, as well as a collection of *Questions*, some of which also bear on our concerns.[64]

(b) Causes and Explanation

As a good Peripatetic, Alexander accepted the doctrine of the four causes (IV.2*a*), applying it to the soul as follows:

382 the soul is both origin and cause for the things ensouled, not only as being productive (*poiētika*) of their vital motions, but also as the cause in respect of form, as was shown, and in addition to these as that for the sake of which, i.e. an end, at least if this completeness of the things which have souls is an end, and what goes before the end is there for the sake of the end. (Alexander, *On the Soul* 24. 11–15 Bruns[65])

end p.356

One of the *Questions* (2. 10) asks how it can be that, given the theory of the four causes, there will not be a material substrate in divine objects; and his endorsement of the doctrine is confirmed at the beginning of *On Fate*: since fate is by common consent agreed to exist and to be a cause of some sort (*On Fate* 2. 165. 15–23), it remains to determine what character fate has (whether it is an ineluctable cause of all events, or merely a predisposing cause of some of them: 2. 165. 23–166. 15), and what sort of cause it is (3. 166. 15–22):

383 the causes of things which come to be are divided into four classifications of cause, as Aristotle showed. For some causes are productive (*poiētika*), some are material in form, while there is also among them the cause in respect of form. In addition to these three causes there is among them the end, for the sake of which what comes to be comes to be. And such is the number of the distinctions of causes, since whatever is a cause of anything will be found to fall under one of these causes. And even if not everything which comes to be requires so many causes, at least those which do require the most will not exceed the number stated. (3. 166. 22–9)

Alexander offers the stock (albeit somewhat misleading) example of the production of a sculpture, in which the bronze is the material, the sculptor the efficient, the particular form ('a discus- or javelin-thrower, or whatever') the formal, and the purpose ('honour to someone or reverence to the gods') the final cause (3. 167. 2–12: cf. Galen: XI.3*a*, **401**). That is, in spite of the intrusion of Stoic terminology (*poiētika*), Aristotelian orthodoxy. Fate must be an efficient cause: the problem is to see what sort of efficient cause, and hence to determine whether it is a complete and ineluctable cause of everything which comes to be (4. 167. 16–19). Alexander first distinguishes between those things which come to be for some end, and those which have no purpose whatsoever (4. 167. 19–26). There is no way to subdivide the latter class; but the former admits of two subspecies, the purposive results of nature, and those of reason (4. 168. 1–18), while there is a third subspecies of things which are fortuitously useful (4. 168. 18–24; Alexander considers the role of luck in greater depth in chapters 7 and 8; cf. *Mantissa* 24). Alexander then argues eliminatively that fate must be 'in the things which come to be by nature' (*On Fate* 6. 169. 128–19).

Here Alexander apparently sees fate as functioning at the level of the general, orderly operations of nature that see to it, for instance, that 'man comes to be from man, and horse from horse' (6. 169. 20–3). Fate corresponds to the natural constitutions of things, in virtue of which, for instance, they are subject to disease and decay, but not ineluctably so, 'for treatments, changes of climate, the prescriptions of doctors, and the

counsel of the gods are sufficient to break down this ordering' (6. 170. 12–16). Thus fate is simply another name for the general dispositions which are the result of the workings of nature. But as any good Aristotelian knows, these produce their effects in the sublunary world at any rate only for the most part (cf. IV.2*b*); and while some things (death and taxes, say) are inevitable, the manner and time of your death, or the amount you owe the Inland Revenue, are not (this is comparable although not identical with the Middle Platonist account sketched above: §3*d*).

Alexander seeks to find a role for fate that will not make it determinative of particular, contingent events: that I shall die is determined by fate, but not how, or when, or where.[66] He is of course particularly concerned to combat the view, paradigmatically associated with the Stoics, that the entire course of the universe is settled from time immemorial (cf. VII.2). In chapter 22, Alexander retails the anti-Epicurean arguments of the Stoics who hold that any introduction of uncaused motion threatens total chaos, and the annihilation of the world's evident order (see VI.2*a*, **226**): consequently, all causes are such as to determine their effects (VII.1*a–b*; **252**). In the succeeding two chapters he attempts to destroy this thesis.

He first notes that some things do not, as a matter of fact, bring about their accustomed effects: although man begets man, some men die childless, yet 'one must posit causes on the basis of likeness' (ibid. 23. 193. 8–12). This last claim (effectively a version of PCS) is used to support the view that some things are causally inert: the causal chain, in this case the genealogical one that links parent and child, simply peters out. Alexander retails the Aristotelian commonplace that many things fail to realize their innate potential (193. 12–14: cf. IV.1*a*).

Alexander is invoking the Aristotelian model of agents as efficient causes. Properly speaking the efficient cause of something is whatever it is that, under the appropriate description, standardly brings about results of this sort; and of course it is, as Alexander alleges, a commonplace that not everything which is fitted by nature to bring about a certain result actually does so. Yet it is unclear how that is supposed to rebut the determinist thesis without simply begging the question: if determinists hold that events, rather than substances, are the proper candidates for causal status, then no reflection on the nature of agent-causation of this sort will serve to rebut their claim. On the other hand, the Stoics were not partisans of event- causation: they too were inclined to see the world as a complicated inter-play

of substances (VII.1*c–g*). But, as they saw, that need not compromise causal determinism.[67]

Alexander in fact appears to confuse the concepts of cause and explanation. Insisting, in orthodox Aristotelian fashion (IV.2*b*), that not every state of affairs can be given an exhaustive explanation in terms of its manifesting the actualization of some substantial potential, he seeks to argue that the world itself is causally disjunct. Thus, seeking to clarify his position, he writes:

> **384** it is possible to save the view that nothing comes to be causelessly even while things are as we state them to be. For if we abandon the causal chain and stop saying that when things have first come to be it follows that they must naturally become causes, as if being causes were included in their nature, but rather make our start in causal ascription from the things which come to be later and further seek those things which are properly causes of their coming to be, neither will any of the things which come to be come to be causelessly, nor for this reason will everything which comes to be be in accordance with such a fate. (ibid. 24. 193. 31–194. 8)

The key, according to Alexander, is to see that causal necessity is a species of hypothetical necessity (IV.3*c*). Sophroniscus' mere existence is not enough to ensure that he will be a father: but if Socrates is to exist, Sophroniscus must be the cause of it (24. 194. 8–15). Furthermore, even accidents have causes, although not causes of such a type as to regularly bring about their effects. Alexander's favourite example is that of accidentally discovering treasure while digging your garden (24. 194. 18–19; cf. 8. 172. 25–31, 174. 14–18; *Mantissa* 24. 177. 35–178. 5):[68] in Alexander's terminology, digging one's potatoes is not a primary (*proēgoumenon*)[69] cause of finding treasure, but rather is merely an incidental cause of it (ibid. 24. 177. 1–179. 6; *On Fate* 8; cf. IV.2*a–b*: **161–2**).

Moreover, Alexander continues (*On Fate* 25), it is a mistake to think that everything which precedes something is its cause. Mere temporal succession is no guarantee of causal consequence; day does not cause night, nor vice versa, 'for in many cases we see that the same thing is the cause of both the earlier and the later occurrences' (25. 195. 5–6; cf. VII.1*f*, **267**). Finally, Alexander claims, it is reasonable to suppose that the causal sequence is not infinite, since the fact that everything generated has a cause does not entail that everything does, since some things are eternal (and hence ungenerated); furthermore, the postulation of an infinite sequence of causes does away with causes (and knowledge) altogether (25.
end p.359

195. 28–196. 7). Again, Alexander's point relates more to the concept of explanation than to that of causation as such; and even if true, it does little to dent the determinist's position. He takes over, defends, and expands Aristotle's perfectly legitimate claims in regard to explanation: but one cannot rebut the position that the causal relation is ubiquitous and transitive by pointing out, however correctly, that the explanatory relation is not.

(c) Fate and Responsibility

In another essay in the *Mantissa* (22),[70] Alexander pursues similar themes in an attempt to delimit an area of action which is genuinely within the compass of free agents, things which are, in a real sense, 'up to us' (*eph'hēmin*). As we saw earlier (VII.2*e*), Alexander spends a good deal of time in *On Fate* arguing that the sense in which the Stoics and their ilk can allow that things are up to us is too attenuated to bear the weight of the moral concepts which are standardly supposed to depend upon it, and without which civilized life will be an impossibility. Equally, their notion of chance as 'a cause obscure to human understanding' (VII.2*a*) radically misunderstands the notion (chances have causes, they just don't have *primary* causes: § *b* above; cf. VII.2*a*, *d*). Alexander's problem is Aristotle's: how can we be said to be free if we are the products of our heredity and environment (cf. IV.4*b*, **177–8**)? If people differ in their natural abilities, while their intrinsic natures plus the habituation they have undergone is responsible for their making the choices they make,

385 one might indeed be utterly at a loss as to how choice could be up to us. For if we make education responsible too, learning will not be up to us either (I mean by 'up to us' that whose opposite is both possible and up to us). And one might be in even worse difficulties if nothing comes to be causelessly: but this is what everyone believed. For the things which are brought about by us in the present must have a prior cause, and the same cause cannot be a cause of opposites; but if this is so, everything which comes to be will do so of necessity, since their causes have been established in advance, at any rate if no uncaused motion is to be found. (Alexander, *Mantissa* 22. 169. 6–170. 7)

The argument makes use of fundamental causal principles we have frequently met in other contexts, namely that nothing comes to be without a cause (CP1*a*: I.3*d*; cf. §3*c* above) and that causes cannot bring about their opposites (P2: III.1*b*). The upshot, if you allow nature and nurture a preeminent role in the development of character, and hence of dispositions to choose, is that without uncaused motions of the Epicurean type it looks as though nothing will be up to us.
end p.360

But that threatens, Alexander alleges, all sorts of dire consequences. What is 'up to us' will consist merely in the expression of our natures, regardless of how they have come to be what they are (*On Fate* 12–14), and that will mean that we cannot be held responsible in any strong sense for our actions (19. 33–8). Indeed, Alexander thinks, it would be better to believe in freedom on the libertarian model even if such a belief is unfounded, since the alternative subverts life altogether (21; cf. VII.2*d–e*; compare Galen's denial of the claim that determinism makes punishment unjust and pointless: XI.5*c*, **431**).[71]

For Alexander, as for Aristotle, the power of deliberation is the mark of both rationality and freedom (11. 14–15): it is this which makes the Stoic idea that liberty simply consists in following nature unacceptable (13). Moreover, deliberation must involve considering at least two conflicting yet equally possible options: Alexander's claim that the opposite of what is up to us must also be up to us (**385**) is emphatically repeated elsewhere:

> **386** as deliberation has been abolished by them, as has been shown, so too has what is up to us evidently been abolished. For this is what everybody (other than those simply defending a paradox) accept as being up to us, namely that over which we have control either to perform or not to perform, not following some external, ambient causes and giving in to them, and following where they lead. . . . Thus choice exists neither in those things which come to be of necessity, nor in those which do not come to be of necessity but not through us either, nor even in all the things which come to be through us, but only in those which come to be through us over which we have control either to perform or not to perform them (12. 180. 3–12)

The libertarian doctrine seems to require uncaused motions; Alexander claims not only that such motions exist, but also that they were noted by Aristotle (*Mantissa* 22. 170. 7–10: the reference is presumably to *Metaphysics* 6. 2–3; cf. IV.2*b*, **161–2**). Yet elsewhere he rejects the Stoic argument that if we allow that someone can react in different ways in precisely the same circumstances, uncaused motions will have to be postulated (*On Fate* 15. 185. 7–11). Clearly, the type of 'uncaused motion' Alexander attributes to Aristotle must not be anything like the Epicurean swerve. Rather, Alexander continues, uncaused motion can be shown to exist 'if there is in some way non-being in things which are' (*Mantissa* 22. 170. 10–16). That opaque doctrine is not much clarified in what follows; but Alexander links the notion of non-being with that of incidental (*per accidens*) being, and remarks that

> **387** whenever what is incidental is present in the same things considered as causes, the cause will be incidental; for whenever anything is consequent upon a

end p.361

certain cause which does not, however, exist for the sake of this consequence, the preceding thing is called an incidental cause, i.e. not a [real] cause. So its consequent came to be causelessly, since it did not do so through a proper cause. (ibid. 22. 171. 10–14)

This is what, allegedly, makes it not the case that we will always be affected in the same way in the same circumstances: the external circumstances themselves are not such as to qualify as proper (i.e. invariable) causes of their particular outcomes. Elsewhere, Alexander prefers to speak of free actions having their causes in the agents alone, rather than being uncaused:

388 if the cause is the origin of those things of which it is the cause, and man is the origin of the things he does, then he will be a cause of them. So if indeed it is absurd to seek and speak of an origin for an origin (for that of which something else is the origin will not then be an origin without qualification), there will not be any pre-established productive cause other than the man's choice, will, and judgement of these things (for then he would no longer remain an origin), but he himself is the cause of the things he does, along with his judgement and choice and their productive causes, but of these things themselves nothing else is <the cause>. (23. 173. 10–17)

But whatever these differences of expression amount to, Alexander consistently rejects the idea that the external circumstances which prompt or inform an agent's action are to be considered as properly the causes of that action. Alexander insists in a number of passages that we are not, as a matter of fact, invariably affected in the same way by the same *external* circumstances, a fact which he takes empirically to refute the Stoic position. We perfectly well can, and on occasion do, he says, choose to act differently on different occasions when precisely the same options present themselves, since we sometimes choose what is noble over what is pleasant and sometimes vice versa (*On Fate* 15. 185. 21–186. 3); sometimes, indeed, we may choose to behave unpredictably simply to refute the determinist (29. 200. 4–7).

We are finally in a position to assess Alexander's doctrine of human freedom, and the extent to which it constitutes a refutation of Stoic determinism. If Alexander is right that precisely equivalent sets of external conditions may yet allow the same agent to behave differently on each occasion, as is plausible, that will indeed show that external circumstances are not *on their own* determinative (at least if similar causes must produce similar effects), which might in turn be thought to refute the Stoic doctrine that fate is an eternal sequence of antecedent causes (VII.2*a–b*; cf. VI.3*b*). But that will only be the case if fate, thus considered, is supposed to be determinative on its own: on the account sketched for Chrysippus in VII.2*a–c*, this need not be the case.

Furthermore, it may well be in principle possible to explain why, in each particular case, the agent's reaction varies even though the surrounding circumstances do not: it will do so because the agent himself will be in distinct, causally efficacious and causally explicable states. And if that is true, the mere fact that external conditions are not the whole cause of an action (and hence cannot properly explain it, which is the proper lesson of Aristotle's *Metaphysics* 6. 2–3, which Alexander invokes above: see IV.2*b*, **161–2**) has no tendency to show that an agent's deliberative choices are not themselves exhaustively conditioned by the entire sum-total of relevant circumstance, including the agent's own evolving states and dispositions. The agent will still be an origin of his action (although perhaps not in a sense strong enough to satisfy Alexander: **388**); but he will not be causally free of the world around him.

5. Conclusions

Plato's successors developed his metaphysical theories, Speusippus concentrating on the overarching One (§1*a*), and Xenocrates on extending the Pythagorean-influenced metaphysics of numbers (§1*b*). Both sought to provide a more complete and comprehensive picture of the intelligibility of the universe. Posidonius opened Stoicism up to influence from both Platonism and Aristotelianism, at the same time necessarily diluting its original purity (§2*a*), and busied himself in un-Stoic fashion with empirical science, particularly astronomy (§2*b*). A generation later, Antiochus sought to lead the Academy away from scepticism and back towards doctrinal Platonism, but Platonism of an equally accommodating, eclectic nature (§2*c*), a move which paved the way for the emergence of a syncretistic Middle Platonism (§§3*a–d*). Against this homogenizing trend, however, a revival of more or less orthodox Aristotelianism was pioneered by Alexander of Aphrodisias (§§4*b–c*), who vigorously attacked Stoic determinism as incompatible with human autonomy and responsibility.
end p.363

XI Science and Explanation

R. J. Hankinson

1. Ptolemy

(a) Saving the Appearances

Plato's challenge to produce a mathematical model of the heavens was taken up by Eudoxus and his like (III.5; V.2*f*, **197**). But it is one thing to produce a predictive algorithm designed purely to model the observable phenomena and quite another to elaborate a concrete picture of the actual structure of the heavens (cf. X.2*b*, **361–2**).[1] There is no doubt, however, that some at least of the astronomers were concerned with the physical realities of their models. Aristotle was not himself an astronomer, but he supplied a (somewhat baroque) physical interpretation for Eudoxus' mathematical model (V.2*f*). Plutarch, in the course of discussing Plato's puzzling description of the earth ‘winding around the pole stretched through the universe’ (*Timaeus* 40b), asks:[2]

388 did he thus assign motion to the earth like that of the sun and the moon and the five planets . . . and should the earth 'winding around the pole stretched through the universe' be considered to have been contrived both to be contained and at rest, but turning and spinning, as Aristarchus and Seleucus later maintained, the former as a hypothesis only, but Seleucus actually asserting it to be true? (Plutarch, *Platonic Questions* 1006c)

Another passage from Plutarch is relevant:

389 Cleanthes thought that the Greeks should indict Aristarchus for impiety for moving the hearth of the universe, because this man tried to save the appearances by hypothesizing a stationary heaven with the earth orbiting along the

ecliptic while rotating about its own axis. (Plutarch, *On the Face in the Moon* 923a)

Thus (at least on Plutarch's account) Aristarchus sought only to save the appearances in the weak sense of providing a mathematical model from which they could be deduced. Seleucus more boldly affirmed heliocentricism to be a physical fact.[3] But Aristarchus' hypothesis won little support, however it was supposed to be taken; for it was swiftly superseded by the latest geocentric (or more accurately geofocal) model of the universe, that supplied by Hipparchus.

Hipparchus (*c*.180–120 bc) was a brilliant observational astronomer. He corrected many of the observations of Eudoxus, as set down in Aratus' third-century poem the *Phenomena*, upon which he wrote a commentary, his only extant work; he discovered the precession of the equinoxes, and calculated its annual effects to within about 10 per cent of modern values; and his estimation of the length of the tropical year was accurate to within six minutes. He was also able to account for the variation in seasonal length noted by Euctemon (III.5) by proposing that the earth was displaced from the centre of the sun's orbit.

But most important from our point of view was his elaboration of a new model of the heavens to replace Eudoxus' concentric spheres (the new model was originally due to the third-century mathematician Apollonius of Perga, best known for his work on conic sections). Hipparchus rather supposed that each planet was carried upon a small circular orbit, or epicycle, about a point (the deferent) which was itself revolving around the earth on a larger circular orbit. The epicycles might revolve in the same sense as the main orbit, in which case it can be shown that the planet will describe retrogressive loops (the number depending directly upon the ratio of the epicycles to the main orbit), or in the opposite direction, which, again depending on the frequency of the epicycles, will produce a variety of smooth curves.[4]

As far as we know, Hipparchus made no attempt to give his system a physical realization. But so successful was it predictively that the new celestial geometry of epicycle, deferent, and eccentric came completely to dominate the subsequent history of ancient astronomy. Indeed, it was dominant for very much longer: the system was refined and extended, as

more accurate and numerous observations were accumulated; but it remained in essence unchallenged until the sixteenth century.

It was this model which was adopted and set out in great detail by Claudius Ptolemy. Ptolemy (*c*. ad 100–*c*.175) lived and worked, as his name might suggest, in Alexandria. He wrote on a wide variety of scientific subjects (his *Optics* largely survives, in a Latin translation), and a short treatise on epistemology and pyschology,[5] as well as composing works on harmonics and geography. Ptolemy's greatest achievement is his *Syntaxis* (often known by its Arabic name, the *Almagest*), a comprehensive, quantitative treatment of astronomy on a modified Hipparchan model which was to be the canonical work on the subject until Copernicus.

The *Syntaxis* is primarily occupied with the task of devising accurate models, involving epicycles and eccentrics, to account for the large body of data by this time accumulated, some of it at least by Ptolemy himself (see *Syntaxis* 5. 1–2 for his new and more accurate observations of the moon in its first and third quarters). He gives detailed descriptions of the instruments he used (1. 12; 5. 1, 12, 14; 7. 4); and there is no real reason to doubt that he made and recorded extensive observations,[6] although it is a further (and vexed) question how these observations relate to the theory he elaborates.[7] It is at least clear that the theory is established on what appears to us at least to be a relatively thin empirical base; and there is no sign of any systematic testing of the theories against new evidence subsequent to their production. Ptolemy wrote an *Optics*; but while he was aware of the distorting effects on observation of refraction (he offers the first detailed account of the phenomenon, tabulating the results for the distinct media of air, water, and glass, and for angles of incidence between 10° and 80° in increments of 10°)[8] , he takes no explicit account of it in his celestial observations (although he will on occasion remark on the difficulties which attend the observation of heavenly bodies, particularly when they are close to the horizon: *Syntaxis* 1. 3, 8. 6, 8. 9. 2).[9]

But whatever their relation with observation, the theories themselves are impressive. He made significant changes in the models (particularly

those for the moon and Mercury), and introduced a new theoretical notion, that of the equant, a point at the same distance from the centre of the eccentric circle of the planet as the earth, and directly opposite it on the same diameter. The deferent of the planetary orbit was judged to orbit the equant with unchanging angular velocity. Thus Ptolemy sought to preserve a uniformity of motion, even though the deferent itself moved with inconstant orbital speed (and, in a sense, anticipated Kepler's third law, and the concept of the conservation of angular momentum).

But although Ptolemy's object in the *Syntaxis* is primarily mathematical, he by no means ignores the physical side of astronomy. Indeed, in discussing the hypotheses of Aristarchus, he writes:

390 some thinkers, although they have no means of refuting the above arguments [sc. for geofocal celestial geography], have elaborated a system which they think more acceptable, and they think that no evidence can be adduced against their hypothetical suggestion that the earth rotates around a single axis from west to east, completing roughly a revolution a day, or alternatively that both heaven and earth have a rotation of a certain amount (whatever it might be) around the same axis,[10] as we said, but such as to preserve their relative positions. These people forget that, while as far as the celestial appearances are concerned there might perhaps be no objection to the simpler version of the theory, yet, judging by conditions as they affect us and in the surrounding atmosphere, this hypothesis must be allowed to be quite absurd. (ibid. 1. 7; cf. 1. 5)

This 'absurdity' is then established by a sequence of arguments which held the day until Galileo's brilliant advocacy of a relativized concept of motion: if the earth moves, it must be moving very rapidly, which motion ought to be detectable in the form of a constant blast of air from the east; moreover, projectiles should move much further and faster westwards than eastwards; but none of these effects is observed. Equally, if the earth orbits the sun, the stars should present different appearances at different times of the year (a fact noted by Aristotle: *On the Heavens* 2. 14. 296^a34–$^b6$); but no such parallax is detectable (it was not in fact detected until 1838). Aristarchus, and following him Copernicus, both assumed, correctly as it turned out, that the stars were so far away as to make such parallax negligible and unobservable (at least for them): but such an explanation has an uncomfortably *ad hoc* feel to it.

Ptolemy allows that the heliocentric system is in a sense simpler than any geofocal alternative; his point is precisely that considerations of elegance should not be allowed to outweigh empirical evidence. **390** also shows him concerned about the physical acceptability of the hypotheses the

astronomer proposes; and indeed without any such realist concerns the distinction between heliocentricity and geocentricity becomes meaningless.[11] And this concern is neither isolated, nor unique to him; Theon of Smyrna, in his astronomical handbook produced a generation earlier, writes:

391 it is not natural for the stars [i.e. the planets] to move themselves along certain circular or spiral lines, and, what is more, against the sense of rotation of the universe, or that they should be borne around in certain circles to which they are attached and which rotate about their own centres. . . . For how could such massive bodies possibly be fixed to incorporeal circles? But there must be spheres, made of the fifth essence, set up and moving in the depth of the universe, some higher, others arranged below them, some bigger, some smaller, and some massive inside the hollow ones, to which the planets are fixed in the same way as the fixed stars. (Theon of Smyrna, *Mathematical Exposition* 178. 12 ff.[12])

Theon is thus well aware of the problem of supplying a physical interpretation for mathematical astronomy. Furthermore, it had been known since Hipparchus that the hypothesis of eccentricity was mathematically equivalent to that involving a particular type of epicycle:[13] and elsewhere, Theon suggests that Hipparchus had preferred epicyclic models to their eccentric equivalents (although he made use of both) on the grounds that it preserved the earth as the focus for the system in a more ‘natural’ manner (ibid. 188. 8 ff.).

Ptolemy too was concerned that the mathematical models be given a physical underpinning; but in view of their growing complexity this was proving increasingly difficult to do. Moreover, this complexity itself could be seen as a defect in the light of an intellectual aesthetic inclined to accept simplicity as a criterion of beauty (a criterion Ptolemy himself adopts, and uses as a regulative methodological principle), something Ptolemy notes and tries to explain away in *Syntaxis* 13. 2.

In a later work surviving largely in Arabic, the *Planetary Hypotheses*, Ptolemy does indeed attempt to supply just such a physical interpretation. After giving a summary of the geometry of the system in the first book, he proceeds in the second to talk of the actual cosmic structure, insisting, in Platonic fashion, that the planets themselves (rather than any body in which they are embedded) must be taken to be the source of their own

end p.368

motion (*Planetary Hypotheses* 2. 12; cf. Plutarch, X.3*c*, **375**). Ptolemy invokes as a methodological principle the Aristotelian notion that nature does nothing in vain: therefore, we must postulate only the minimum of physical apparatus required to produce the phenomena (*Planetary Hypotheses* 2. 6). Thus the planet is conceived as being mounted upon a wheel rather than a sphere, this wheel (the physical epicycle) being rotated between two other concentric wheels with the earth lying either in or slightly off their centres, as the particular model demands. Each of these systems is independent, and hence Aristotle's countervailing spheres are also done away with (2. 6). This independence is underlined by a vivid simile:

> **392** let us choose to illustrate the motions of the heavenly bodies with the well- known movements of birds. . . . The source of their movements is in their vital force which produces an impulse that ramifies into the muscles and thence into the feet or wings. . . . There is no compelling reason at all to suppose that the motions of all the birds occur through mutual contact. . . . Equally, we must suppose that of the heavenly bodies each planet possesses its own vital force and moves itself and causes motion to those bodies which are naturally conjoined with it. (2. 7)

The single, unified, cumbersome system of Aristotle is replaced by the image of a flock of planets all flying like ideal birds in perfect formation, or, to adapt the resonant metaphor of the *Epinomis* (982d–e; cf. *Timaeus* 40c; III.5), dancers performing a perfect minuet (*Planetary Hypotheses* 2. 8; cf. ps.-Aristotle, *On the World* 6. 399^{a}11–26; Proclus: XII.3*b*).

(b) Astrological Explanation

Like many famous figures in the history of astronomy, Ptolemy sought to extend the range of his subject's predictive powers beyond the mere structure of the celestial revolutions and into human affairs. His presentation and defence of astrological prediction, the *Tetrabiblos*, survives, and is a significant contribution to a debate we have

already considered (VIII.5*b*).[14] Ptolemy is well aware that astrology, unlike astronomy proper, is not an exact science (*Tetrabiblos* 1. 1, p. 1 c[15]), and that this has caused it to be disparaged. But this uncertainty is only to be expected in a science involving the intrinsic unpredictability of matter (1. 1, pp. 1–2 c). Even so, Ptolemy concedes that the objections have some force on their side, and thus undertakes to defend both the possibility and the usefulness of astrology against its detractors (1. 1, p. 2 c).
end p.369

Ptolemy takes over from Aristotle the five-element cosmology, with the crucial Stoic addition of the notion that celestial events can indeed have an influence upon those in the terrestrial sphere (1. 2, pp. 2–3 c). This is of course obvious in the case of the sun, a fact which Aristotle also naturally recognized (V.2*g*, **198**); but equally, the moon too can be shown to affect tides (cf. X.2*a*) and the life-cycles and behaviour of animals (1. 2, p. 3 c). Ptolemy then continues:

393 the passages of the stars too (both fixed and planetary) produce many indications of torrid, windy, and snowy atmospheric conditions, by which things on the earth are affected appropriately. Moreover, their aspectual relations with one another, by the meeting and mixture of their influences, contrive many subtle changes; for even though the sun's power is paramount with regard to the ordering of quality in general, the others collaborate with or frustrate it in particular ways, the moon more obviously and continuously in being new, half, or full, the other stars less frequently and obviously in their risings, settings, and conjunctions. (1. 2, pp. 3–4 c)

The cautious tone in all of this is evident; the argument is a kind of induction, moving from the evident case of solar power by way of less obvious lunar effects to the hypothesis of other astral influences. Note also the move in the first two sentences from astrology as merely indicative ('soft' astrology: cf. VIII.5*a, b*), to the claim of actual celestial influence upon earthly events. In what follows, Ptolemy notes that sailors, farmers, and other professionals make use of astronomical indications to warn them of storms, and to determine the right time for planting (1. 2, pp. 4–5 c); but that fact alone will not support the hypothesis that the celestial bodies *affect* the seasons, as opposed to simply marking them.

Ptolemy readily concedes that astrologers frequently err through ignorance and poor calculation, as well as because of the sheer difficulty of the art (1. 2, pp. 5–7, 9 c); but that is no fault of the art itself (compare *Science of Medicine*: II.2*d*, **104–9**). Properly conducted, he thinks, there is no reason why astrology should not be able to discern the general shape of individuals' characters from their horoscope (1. 2, pp. 5–6 c). But in an important concession, which goes a long way towards undermining the empirical anti-astrological arguments of Cicero, Sextus, and others (VIII.5*b*), he is careful to insist that astrology on its own cannot predict particular, individual events, nor can it precisely determine the future course of an individual's life, since it is not the only force that operates in such affairs:

394 so concerning the investigation of what goes in the surrounding heavens, this [i.e. the difficulty of making precise enough observations] would be the only impediment, since in this case there is no cause to be considered apart from the motion of the heavenly bodies. But concerning the casting of horoscopes,

end p.370

and in general the individual's particular temperament, one can see that circumstances neither small nor insignificant are co-operative causes (*sunaitia*) of the idiosyncrasies of those who are born. (1. 2, pp. 7–8 c)

Ptolemy goes on to specify these *sunaitia*: the species to which each belongs (determined by its specific seeds), the location of its birth, and the way in which it is brought up and the prevailing customs all contribute to the individual's character (1. 2, p. 8 c; cf. II.1*c*, 3*a*); and

395 unless each of them is taken into account along with the causes derived from the surrounding heavens, even if this possesses the greatest power (for it is a co-operative cause of their being the way they are, but not vice versa), they can be the source of much difficulty for those who think that in such cases everything, even those things not wholly under its control, can be discovered from the motion of the heavenly bodies alone. (1. 2, pp. 8–9 c)

Ptolemy concludes by comparing the case with medicine: no doctor is censured for inquiring into his patient's background before making a diagnosis; neither should astrologers be denigrated for speaking of ethnicity, habitation, education, and so on (1. 2, p. 10 c).

In the succeeding chapter, Ptolemy tries to establish that astrology is also beneficial, against the arguments of people like Favorinus (VIII.5*b*) to the effect that it is either superfluous (since if the future is fixed there's nothing you can do about it), and perhaps actually harmful (since it induces pointless anxiety). Ptolemy rejects the latter claim out of hand, on the basis of his objections to the former. Since his astrology is not deterministic, it functions in exactly the same way as medicine: the physiologist can tell you what sort of ailments you are prone to on the basis of your temperaments, and hence prescribe a regimen whereby you may avoid them; so too astrology can provide the foresight with which one may avoid disaster (1. 3, pp. 9–11 c). Moreover,

396 we should not think that all human affairs derive from the heavenly cause, as if they had been ordained from the beginning by some irrevocable divine edict, and fated to occur of necessity with no other cause capable of preventing them. (1. 3, p. 11 c)

The heavens supply the primary causes of things, but not their total explanation; and the primary causes may on occasion be overridden by others, in just the same way as medical intervention can prevent the onset of disease (1. 3, pp. 11–13 c); the astrologer thus deals in defeasible conditionals (if nothing is done, something dreadful will happen), rather than adamantine categoricals (1. 3, pp. 13–17 c; compare Plutarch and pseudo-Plutarch on fate: X.2*d*, **376–7**).

It is obvious that this sophisticated position is far more resistant to the sceptical attacks upon it than those more naïve and far-reaching claims

end p.371

attacked by Cicero and Sextus (VIII.5*b*), and apparently championed by proponents of a more naïve divination such as Manilius.[16] Ptolemy explicitly rejects astrological determinism (1. 3, p. 15 c), and his open acceptance of other causal influences on people's lives affords him the means to evade the twins argument and the problem of natural disasters killing multitudes with widely differing horoscopes (1. 3, pp. 11–12 c; cf. VIII.5*b*).

The rest of the *Tetrabiblos* is devoted to an exposition of the types of astrological influence Ptolemy countenances. It is noteworthy that, after the general treatment of planetary effects in the remainder of book 1, he deals first with general ethnic and natural characters in book 2 as being more important (2. 1, pp. 53–5 c; cf. 1. 3, pp. 11–12 c) before turning to individual horoscopes in the rest of the treatise. This is still hard astrology: in the introduction to the individual treatment in book 3, Ptolemy stresses that the stars are causes (3. 1, pp. 103–4 c); but this is hard astrology shorn of its most obviously absurd claims. In this context it is worth stressing Ptolemy's implicit rebuttal of the problems raised by Sextus concerning the precise establishment of times (VIII.5*b*). Ptolemy agrees that the time of conception is of fundamental importance, but argues that since the time of birth is largely conditioned by the astrological aspect at conception it can still function as a reliable sign of a direct astral causal relationship even though not exhibiting one in itself (3. 1, pp. 105–6 c).

Ptolemy is relatively successful, then, in answering the more pressing theoretical and empirical objections of the sceptics to the possibility and practice of astrology. But he himself provides little in the way of what might count as empirical support for his own claims (e.g. that Saturn's basic powers are to cool and dry moderately—1. 4, pp. 17–18 c—or that Venus warms moderately but humidifies: 1. 4, p. 18 c), although the planetary powers are related (rather feebly) with their relative proximity to the sun. Even further removed from any evidence are the suppositions regarding the relative beneficence or maleficence of the planetary influences, resting as they do on the view that the hot and the wet are ‘fertile and active’, while the cold and dry are destructive and passive (1. 14, p. 19; the rest of the structure is even more fanciful). In so far as Ptolemy immunizes astrology from direct empirical refutation of the sort practised by Cicero and Sextus (VIII.5*a*–*b*), he also dilutes its empirical content to the same extent, rendering its claims ever more remote from anything approaching experimental confirmation or disconfirmation.

end p.372

2. Galen: The Structure of Causation

(a) Background

Galen (ad 129–*c*.215) was the scion of a wealthy family in the rich Greek city of Pergamum in Asia Minor. His philosophical education was liberally broad, comprehending Aristotelian and Stoic logic as well as the prevailing Platonism of the times; moreover, he was familiar with, if not friendly towards, Epicureanism.[17] When he turned to medicine, he travelled to Smyrna, Corinth, and Alexandria in search of the best available teachers. His intellectual outlook reflected this catholicity of upbringing. He composed a major treatise (*On the Doctrines of Hippocrates and Plato*: 5. 181–805[18])

designed to exhibit what he took to be the fundamental agreement on all important questions between his two great masters, Hippocrates in medicine and Plato in philosophy, although Galen lards his Platonism with a great deal of Aristotelian doctrine. By contrast with his favourable, not to say hagiographic, attitude to the original Masters, Galen is not uniformly pleasant about the early Stoics (Chrysippus' account of the emotions and of the structure of the soul is the subject of a particularly virulent attack in *On the Doctrines of Hippocrates and Plato*), while both Herophilus and Erasistratus, although praised in some contexts, also come in for some bilious criticism. But he reserves the bulk of his vitriol for the Atomists, and their medical descendants the Asclepiadeans and the Methodists. Their greatest sin (among many) in Galen's eyes is their refusal to acknowledge the evidently purposive structure of the universe. Indeed, it is for his failings as a teleologist as well as for his denial of antecedent causes that Galen takes Erasistratus to task (*On the Natural Faculties* 2, *passim*; cf. IX.1*c*). Erasistratus did not reject teleological explanation out of hand (Galen admits that he allows Nature some explanatory role: ibid. 2. 83, 88, 91): but he severely restricted it, holding, for instance, that the spleen has no function (2. 91–2; *On the Function of the Parts* 3. 315–16[19]). We shall examine Galen's teleology more fully later on (§3). But let us begin with his attempt to refurbish the concept of an antecedent cause against Erasistratus' onslaughts.
end p.373

(b) Erasistratus and Antecedent Causes

Erasistratus tried to undermine the concept of antecedent causes altogether (IX.1*c*): the things alleged to be the initial causes of illness, such as heat, cold, sleeplessness, excessive eating and drinking, grief, indigestion, anger, overwork, poisons, venoms, dog-bites, and the like,[20] are not invariably correlated with the diseases which they are supposed to cause. Of a thousand people in a theatre on a hot summer's day (and hence all exposed to the same causal influences), only four develop a fever, and only one of these becomes seriously ill (*On Antecedent Causes* 2. 11, 10. 126[21]). None of these factors is sufficient for its supposed effects: and so, Erasistratus argues, they cannot be causes. We have already noted Celsus' answer (IX.1*c*, **325**): these factors, even if not sufficient, seem at least generally to be necessary conditions of illness; and it is possible to give an account of the differing structures of individuals' bodies which will account for the fact that only a few people suffer when exposed to the same causal influ- ence (IX.1*c*, **322**).

Galen, in developing his own account of causation, follows Celsus in urging that causes need not be sufficient for their effects. The antecedent causes that set in train the events ultimately leading to the illness and the standing conditions of the body that render it liable to such affection both contribute to the genesis of disease:

397 for surely if they consider that causes of some kind exist, then they must either say that they require material on which to act, or they must hold that material is impassive, and cannot be affected.[22] If impassive, it is clear that they cannot preserve material itself: for material has its existence in its capacity to be acted upon and readily affected by an agent. And if they concede that the material is acted upon, then in no way will they be able to deny that this happens sometimes more and sometimes less, at least if they concede also that the material is sometimes stronger and sometimes weaker than the cause acting upon it. But if they do concede all these things, what absurdity would there be in saying this: 'Because only four out of the thousand present at the theatre developed fevers, it follows that this four alone were burned'? For if, instead of four, only one had developed fever, there would be nothing to prevent him alone of all of them from being harmed; for what is absurd about him being the only one in a suitable condition to be harmed by

this affection in this period of time, while some of the rest were able to disregard the heat on account of their bodily health, whereas others, although weaker with regard to the heat, would none the less have needed to be exposed to it for a long time before the harming effect culminated in a fever? (ibid. 8. 98–101)[23]

Materials are more than merely the location of change. The type of material in question, and crucially its resistive powers, will determine the extent to which (if at all) the external causal influence produces a perceptible effect in it. Earlier (1. 7), Galen suggests that Hippocrates uses the term *prophasis* for antecedent causes (see II.1*b*) because antecedent causes make manifest (*pro-phainein*) pre-existing weaknesses in the bodies upon which they operate. By insisting that causing is co-operative, Galen can dispel the initial plausibility in Erasistratus' claim that genuine causes must always produce their effects (cf. also VIII.2*c*, **302**). Perhaps to give a full account of the genesis of some particular disease one must spell out all the relevant factors, of which the antecedent cause is only one among several (the duration of exposure to the triggering agent, the patient's disposition, etc.). But that has no tendency to show that antecedent causes are not causes at all, or that they have no causal relevance.

Duration is important. Galen insists that, while individuals vary in their initial resistance to particular types of external influence, the strongest constitution can ultimately be broken down, either by the power of the affecting cause, or by some unfortunate concatenation of circumstances, or because of the length of time to which the body has been subjected to it. Here Galen's humoral pathology enters the picture. Galen adopted the developed Hippocratic theory of temperaments, the blending in the body of hot, cold, wet, and dry (II.1*c*–*d*, **86**, **93–4**). According to your natural constitution, the extent to which you are hot and dry rather than cold and wet (let us say), the skilled medical practitioner can discern the types of causal influence to which you are peculiarly susceptible. For instance, as someone with a natural preponderance of hot and dry, you will be particularly prone to ephemeral fevers brought on by exposure to heat.[24] Thus Galen offers not only a general theoretical account of how antecedent causes may fail on occasion to produce their expected effects; he also deploys a physiological theory of

some complexity that enables him to account (at least to his own satisfaction) for the observed differences between individuals' reactions to the same, or similar, influences.
end p.375

And for Galen, as for the majority of the ancients, these causal relations were universal (cf. II.1*b*, **72**). In a passage of central methodological importance from *On the Therapeutic Method* 10. 36–7, he treats as a priori certain axioms, on a par with the Euclidian axiom that equal quantities substracted from equal quantities leave equal remainders, and the law of the excluded middle,[25] the following causal principles:

(GP1) nothing occurs without a cause (CP1*a*: I.3*d*);

(GP2) everything comes to be from something which exists (CP1*b*);

(GP3) nothing comes to be from the absolutely non-existent (which is a corollary to GP2, and a version of CP1*b*);

and

(GP4) nothing is annihilated into the absolutely non-existent (CP2:I.3*e*).

Of these four, GP1 in particular is a constant refrain throughout his writings (see §5*b* below). In *On the Doctrines of Hippocrates and Plato* (5. 566–7), he adds two further axioms:

(GP5) what is changed takes on a form similar to that which causes the change (expressed slightly differently at *On Semen* 4. 556, 563; cf. PCS: I.3*c*; III.1*b*, 4*a*; X.4*b*; compare Plotinus' PPA: XII.2*b*);

and

(GP6) it is impossible that, when two bodies come together, they should not both act and be acted upon (cf. *On the Natural Faculties* 2. 161).

GP6 amounts to an explicit acceptance of the consequence that Sextus felt to be fatal to the Dogmatists' conception of causation (VIII.3*c*, **302**), namely that causing is co-operative, and that agents as well as patients are affected in the course of causal interaction. The elements of Galen's system of causation, as well as the metaphysical theses that underpin it, are now in place.

(c) Antecedent, Preceding, and Containing Causes

Antecedent causes matter from a practical as well as from a theoretical standpoint, as the rabid-dog case from *On Sects* demonstrates (IX.3*b*). The terminology of antecedent and containing causes is Stoic in origin; and
end p.376

earlier (VII.1*c*–*d*) we reviewed the evidence for attributing to the early Stoics a tripartite division of causes, including the third category of preceding causes (*aitia proēgoumena*). However, a passage from Galen's *On Containing Causes* apparently attributes the distinction (at least in its medical form) to Athenaeus of Attaleia, the first century bc founder of the Stoic-influenced[26] pneumatist school of medicine:

398 the distinctions [sc. in causal type], which Athenaeus held to be three, are these: first there are containing causes, second preceding causes, and third antecedent or procatarctic causes. They call the latter everything which exists outside the body and harms it, bringing on illness, while those which are of the kind that work within the body are called preceding causes, while the alterations of the innate *pneuma* which are brought about by them and even by externals such as the moistening, drying, cooling, and heating of the body he calls containing causes of diseases, since this *pneuma* permeates throughout the uniform bodies, and alters them as it alters itself. (Galen, *On Containing Causes* 2. 2–3[27])

The scheme thus attributed to Athenaeus[28] is essentially adopted by Galen (although purged of its pneumatist implications). This passage is particularly important in that, while emphasizing the Stoic origin of the term *aition sunektikon*, it supports the claim Galen has made in the preceding chapter that, in its original Stoic usage, a containing cause was a cause of being, not of becoming (cf. VII.1*b*, **255**).

By contrast, Athenaeus (and following him Galen) reserve the term for the causes of events or actions, not for the persistence of states of affairs. Elsewhere, Galen writes:

399 so that we will not appear to be disputing about words, we agree to call some causes thus [i.e. *aitia sunektika*] but not, by God, of things that simply exist, but of things whose nature it is to come to be. (*Against Julian* 18a . 280)

Thus for Galen containing causes are present causes of present effects: to take a typical example[29] the tightening of the choroid membrane is the containing cause of the looseness of the pupil of the eye. This suggests that containing causes must also be functionally related to their effects, a suggestion borne out by pseudo-Galen, *Medical Definitions* 19. 393: *aitia*

end p.377

sunektika are causes such that ‘when present the effect is present, when absent the effect is absent, when increased the effect is increased’.

Antecedent causes now fit neatly into the scheme as the external factors that set a primed dispositional mechanism in motion (cf. VII.1*d*, 2*c*). The status of *aitia proēgoumena* is, however, less clear (cf. VII.1*d–e*, **262–3**). Galen sometimes employs the term (or cognates) in reference to *any* causal factor that precedes the effect in question, whether internal or external.[30] But where he does discriminate, he adheres to Athenaeus' distinctions; and he remarks non-commitally that you could do worse than follow the example of ‘some doctors’ in distinguishing antecedent from preceding causes (*On the Causes of Diseases* 7. 10).

At *On the Therapeutic Method* 10. 65–7, Galen distinguishes between excessive ingestion of food, which is the external, antecedent cause of illness, and the excess of blood in the veins caused by it (the *plēthōra*: see IX.1*c*), which is its preceding cause.[31] Here is Galen's fullest exposition of the doctrine, employing a slightly different example:

400 speaking generally, things which are external to a body and alter it in some way are called antecedent causes, because they precede the dispositions of the body. Whenever these dispositions condition containing causes, they are preceding causes of them. For instance, external cold brings about constriction of the skin, and as a result of that constriction normal exhalations are checked, which, being checked, form a collection, causing a fever to take hold, which alters the function of the pulse, which in turn changes the pulse itself. In this case the antecedent cause is the external cold; everything else up to the alteration of the function of the pulse are preceding causes; and through the mediation of the preceding causes, the antecedent cause alters the function of the pulse, which is one of the containing causes, and this in turn brings about a change in the pulse. (Galen, *On the Causes of Pulses* 9. 2–3)

These 'dispositions' are not the standing conditions of the body prior to the action of the antecedent cause; nor are they general susceptibilities to disease (although Galen explicitly believes in such susceptibilities, and thinks he can give a causal account of them). Rather they are the conditions that obtain *after* the action of the antecedent cause, when the mechanism is already operating (this fact distinguishes them from the 'causae principales' of Chrysippus, if those are to be interpreted as a separate Chrysippean category: VII.1*d–e*: **262**). These may be several in number, forming a determinate causal chain: but they are not essential to all causing. What is necessary is the concatenation of internal and external features in the genesis of the disease, a concatenation the elucidation of which both enables Galen to rehabilitate the external factors as genuine, if

causally insufficient, causal factors, and which underpins his classification of physiological types on the basis of differing susceptibilities to these external influences.[32]

3. The Aristotelian Influence

(a) Four Causes?

Thus Galen adopts and adapts the original Stoic causal categories in order to refute Erasistratean causal scepticism in *On Antecedent Causes*. But he also contrives in the same text to offer an account of explanation which is clearly Aristotelian in inspiration (cf. Alexander: X.4*b*):

401 (i) in no human art (*technē*), then, does any result ever come about without a combination of causes. The carpenter would never make a bed unless there were a need for one; and if one were needed, but he hadn't got a saw, an axe, a gimlet, a rule or the rest of the carpenter's tool-kit (or if he had them, but had no wood), he would not be able to construct a bed. (ii) Similarly, if tools and wood are available, and there is a considerable need for its construction, yet the craftsman is missing, then there is no way that the bed can come to be. And the same holds good for all other arts: for they all work towards some end, and they all require material of some sort, and instruments. (iii) Those things that come to be in a different manner, by chance of some sort, do so when one of the types of cause is absent (namely that because of which). For that is what a chance occurrence is: something which occurs for the sake of no end. (iv) Sometimes too the things with which things come to be are absent (i.e. the instruments). (v) And it is clear that we have omitted nothing: that because of which the thing that comes to be comes to be is the need or purpose of its generation; that with which, the instrument; that by which, the artisan; and that out of which, the material. (vi) And it is clear that there is nothing that simply does not require material from which and an efficient cause by which it comes to be . . . it always requires the conjunction of something which is acted upon (by way of being the material) and something that acts upon it (by way of being the agent), sometimes also requiring some instrument, and acting for the sake of something, sometimes without instruments, and not for the sake of anything. (Galen, *On Antecedent Causes* 7. 68–72; cf. 6. 55–67; **402**)

That passage is Aristotelian in general tone (cf. IV.2*a*), but with the striking development that the formal cause has been replaced by the instrumental cause we have already met among the Middle Platonists (X.3*b*, **368–9**; cf. X.1*b*, **354**; 2*c*, **364**(ii)).
end p.379

(b) The Instrumental Cause

The instrumental cause is non-Aristotelian. Aristotle does refer to instruments (‘tools and drugs’) in his causal typology (*Physics* 2. 3. 195^{a}1; cf. IV.2*a*; compare Alexander, X.4*b*, **383**), but not as independent parts of the taxonomic scheme (although Simplicius interprets the passage thus: XII.3d, **487–9**). Galen, following Philo's lead (X.3*b*, **368–9**), elevates them to that position: **401**(i) and (iv) indicate that instrumental causes are to be treated on the same level as the rest of the Aristotelian canon. Indeed when Galen works out his own (non-Aristotelian) version of the distinction between *per se* and incidental causes, he explicitly classes instrumental causes among the genuine causes of things:

402 (i) all men call those things which of their own nature contribute to something's coming to be 'causes'. You can call them 'causes from their own nature', or 'causes properly so called', or 'causes in virtue of themselves': it makes no difference. (ii) The incidental is opposed to all of these cases. The carpenter, the gimlet, and the wood all contribute from their own nature to the creation of the storage vessel: for you can point both to the wood and to the holes bored by the carpenter with the gimlet in the actual thing constructed. (iii) But the location where the wood which the carpenter puts together was worked on, and the surrounding air, will be called causes only incidentally, and not in a primary sense: because if the craftsman is a man, then necessarily he occupies some place. (iv) And if this is so, and it is also necessary at all events that there is no solid body between him and the wood (since that would prevent him from exercising an effect upon it), then it is clear that the intervening space must be either entirely empty, or if not empty, then at any rate occupied by some yielding body; from which it is clear that even the air itself contributes to the creation of the storage vessel. (v) You will understand this more clearly if you imagine the craftsman permeating the whole object, as its nature permeates the whole of it, and not just touching its surface as we do: in those circumstances, it is clear that the creator will not require a location, other than that of the material itself, nor either air or empty space around it. (*On Antecedent Causes*, 7. 76–81)

Even so, he will on occasion maintain that the efficient and the final are the most important of the causal categories (e.g. 6. 67), a preference reflected throughout later Platonism (cf. X.1*b*, **354**, 3*b*, **368**; XII.3*a*, **470**). And in a passage of *On the Function of the Parts*, in which he attacks Asclepiades for mistaking the proper explanation of the structure and function of the 'arterial vein',[33] Galen makes the final cause primary (*On the Function of*

the Parts 3. 464, 470), sometimes pairing it with the material cause (3. 466) and sometimes with the efficient (3. 471–2).

Asclepiades, according to Galen, reverses the proper direction of explanation (the 'arteries' are not thin because they work hard, but are thin in order to work hard), and as a result fails to appreciate the fact that the Creator so made them because they are better that way, which is the fundamental explanation for their condition (on the anti-teleological cast of the Asclepiadeans, compare *On the Function of the Parts* 3. 364). The thinness of the arteries (and their consequent vigorous motion) are merely the instrumental causes with which the Creator effects his design, although even so they are causes, and not mere prerequisites (**402**(iii–v)). Thus Asclepiades

403 overlooks these two causes, the one from the Creator's providence which is said to be primary, and the other as it were the material, and arrives at the most insignificant type of cause of all, one which nobody versed in logic would, I think, call a cause without qualification, but rather an incidental or consequential cause, like a counterfeit drachma. (ibid. 3. 466; cf. **407**).

Yet vigorous vascular motion is not an obvious candidate for causal instrumentality, nor is it on all fours with Galen's examples of **401**(i) (although he allows that a leg may be the instrumental cause of a trampling: *On Antecedent Causes* 7. 73). Vigorous motion is a process or activity, and as such seems if anything to be an efficient cause. But this will cause Galen no trouble. How a particular part of a causal sequence is to be classified depends upon the perspective from which it is viewed. A muscle's action may be the instrumental cause of motion when considered in relation to that motion's proper origin (namely the regent part of the soul), but an efficient cause from the point of view of the bones which are moved (*On the Function of the Parts* 4. 347). Equally, final causes may themselves have further final causes (*On Antecedent Causes* 6. 58–9): if you ask a weaver why he is weaving, he may reply 'in order to make a cloak'; but this admits of a further question, to which the answer will be 'for covering', which in turn is a protection against the cold; and so on (compare Plato: III.3*b*, **133**).

Things may thus be instrumental causes in two different ways. They may actually be tools, items used in the construction of something which contribute essentially to its nature (I shall take this up in a moment); or they may be the intermediate stages in a causal sequence by which some remote result is brought about, the *sunaitia* of Plato's Demiurge, who is so much at the forefront of Galen's mind here (Cf. III.4*b*, **144**; compare Philo, X.3*b*, **369**; and Simplicius, XII.3*d*, **486**).

end p.381

(c) Incidental Causes

403 suggests that instrumental causes are only incidental; yet this is not Galen's view elsewhere. Indeed, in *On Antecedent Causes* he explicitly contrasts instrumental causes (along with the genuine Aristotelian triad) with incidental causes:

404 in order to act upon the wood, the carpenter needs an unimpeded intervening space, not in so far as he is a carpenter, but in so far as he has a body which occupies some place other than that of the material being worked on. These things, then, are incidental, and have the status of prerequisites, cutting off and preventing a thing's creation if they are not present, but contributing nothing essentially by their presence. In fact you should not say that they cut off the effecting but rather its essential origins, when they are not present. (*On Antecedent Causes* 7. 83–4; cf. **402–3**)

The terminological differences between **403** and **404** are trivial and easily explained away (in **404** Galen seeks to develop a semi-technical meaning for the term 'incidental cause'; in **403** he merely wishes thereby to downplay the importance of instrumental causes; compare **407** below). But why are instrumental causes genuine (if relatively minor) causes? Alternatively, why should such factors as the 'unimpeded intervening space' (**402**(iii–iv); **404**) that allows the craftsman to produce his artefact be reduced to the status of a mere *sine qua non*?

The eclectic Potamo of Alexandria (fl. *c*.25 bc) did indeed treat place ('the in which') as a genuine causal category (Diogenes, *Lives* 1. 21; cf. Iamblichus, XII.3*a*, **470**, for another sense in which place may be a cause). In rejecting this, Galen suggests that a distinction might be made between the contributions made by place (and unimpeded intervening space) on the one hand, and tools on the other, to an outcome. The carpenter must occupy some place if he is to produce something: but there is no particular place he must occupy (compare Clement on the status of time: VII.1*f*, **265**). Equally, the type of intervening

fluid makes no difference to the outcome, even if it might make it easier or more difficult of realization (a carpenter could work underwater if he were sufficiently perverse).
By contrast, it seems reasonable to think that the type of tools involved will directly affect the outcome. Just as one can explain some of a statue's properties at least in part by reference to the material from which it was made, equally the type of tool employed may make a genuine and non- incidental contribution to the outcome: the silky finish is due to the fine- grained sandpaper. By contrast, incidentals such as time and place are in a sense necessary conditions of the result (**402**(iv)), but not 'from their own nature' (**402**(i–ii); **404**).
end p.382

Somewhat confusingly, there is yet another sense in which causes may be incidental for Galen. Here the issue is not whether they genuinely contribute anything from their own nature to the effect in question, but rather their proximity or remoteness from it. In *On the Differences of Symptoms* Galen invokes once more his sub-Aristotelian causal scheme:

405 causes are many in genus: there is the material, the purpose (*chreia*), the goal (*skopos*), the instrument, and that which initiates the motion. (*On the Differences of Symptoms* 7. 47–8)

This division is then applied to causal sequences. In any causal chain, the links are related to each other in a variety of different ways; but Galen chooses to concentrate, for the purposes of this passage at least, on their relative proximity. Take any causal sequence of events, for instance a series of falling dominoes, arranged such that the first causes the second to fall, the second causes the third to fall, the third the fourth, and so on. In the language that Galen introduces here, the first causes the second to fall 'primarily and essentially', but causes the third to fall 'secondarily and incidentally' and it causes the fourth to fall 'thirdly'; and so on; while the second causes the third to fall 'primarily and essentially', but causes the fourth to fall only 'secondarily and incidentally'.
In any causal sequence, each event is linked to its immediate successor in a relation of primary and essential causing, but to its more remote descendants as an incidental cause. The remoteness of a cause from the effect in question does not entail, *contra* Erasistratus, that it cannot be a cause of it at all: the causal relation is transitive, and Galen's reply to Erasistratus holds. But for all that, such perspectival distinctions may serve useful purposes.
To call a cause 'primary and non-incidental' in this sense does not confer on it some particular objective status: it is so only in relation to some other individual. Congruently, a cause is not 'primary' (in this sense) if it is the most important cause in the sequence considered from some independent standpoint. This categorization has nothing to do with objective assessments of differing causal contributions.
One further point should be made in connection with **405**. Galen distinguishes between the function or purpose (*chreia*) and the goal (*skopos*) of a particular activity. The Stoics differentiate between the *telos*, or end, of some activity or sequence of events on the one hand, and the *skopos*, or goal, on the other.[34] Galen (who sometimes uses the terms interchangeably: *On Antecedent Causes* 6. 57), however, distinguishes between them as follows:
end p.383

406 the *skopos* of the art of medicine is health, but its end is the possession of health. (Galen, *On Sects* 1, 1 H)

You direct your attention towards the *skopos*; the end is some state of affairs associated with that thing. The Stoics considered that living rightly consisted precisely in organizing one's activities towards the proper *skopoi*, the actual achievement of them being unimportant to happiness.[35] Galen does not adopt that somewhat paradoxical view; but he will still on occasion find a use for the distinction between what it is that organizes our activities, and those activities' actual ends.

(d) The Formal Cause

We saw how earlier syncretizing philosophers were prepared to adopt Aristotle's scheme with the addition of a transcendental paradigmatic cause answering to Plato's Forms (X.1*b*, **355**; 2*d*, **366**; 3*d*, **379**; cf. XII.3*d*, **486**). But Galen has surprisingly little to say about form in either its Platonic or its Aristotelian guise. Indeed he mentions the formal cause only when attacking Asclepiades' account of the structure of the pulmonary 'artery' (§ *b* above), in a passage worth quoting at length:

407 the first cause for everything which comes to be is the aim (*skopos*) of the activity, as Plato has indicated somewhere.[36] Thus, when asked the reason for coming to the market, one cannot omit this reason and give a better reply in some other way. It would be absurd for someone were rather to say that he had come to the market because he had two feet which could move him easily and support him safely on the ground, instead of saying that he had come to buy provisions or a slave, or to meet a friend, or to sell something. He might perhaps have given some reason, but not the real or primary cause, rather something instrumental, indeed more of a prerequisite than a cause. Plato correctly understood this about the nature of causes [cf. III. 1*a*, **123**]. However, we will agree that there are more types of causes in order to avoid verbal quibbling: the first and most important (*a*) that because of which something comes to be; the second (*b*) that by which it comes to be; third (*c*) that from which; fourth (*d*) that by means of which; and fifth, if you want, (*e*) that in accordance with which. (Galen, *On the Function of the Parts* 3. 465)

(*a–d*) are the sub-Aristotelian categories we have already examined; but (*e*), the formal cause, is added more or less as an afterthought. Having set Asclepiades straight in regard to 'arterial' structure, he adds:

408 you have in this account all of the causes: that derived from the end, that from the Demiurge, that from the instruments, that from the material, and that in accordance with the form. (3. 471)

end p.384

But he does nothing whatsoever with the fifth category; and this passage is unique in Galen's work in referring to the formal cause at all.

We have already seen that the Stoics' containing cause was (originally at least) 'the Stoic analogue to the Aristotelian form';[37] and Galen refuses to allow that there is any need for such a cause of being (VII.3*b*, **255**). Complex objects may sometimes have physical reasons for their cohesiveness (pegs, ligaments, and so on), but the whole point of postulating elements is that they resist any further attempt at explanation: causes of being are vacuous. If Galen took Aristotle's form, then, simply as being the essence of the particular natural kind (as opposed to being a dynamic internal source of coming to fruition), he may well have thought it served no useful explanatory purpose, at least as a cause. Other passages show how alive Galen is to the dangers of such pseudo-explanation:

409 there are those who say that the heating of the head by the sun is the containing cause of the warmth that results in it, and the cut the containing cause of the wound. But they do so only because they do not realize that they are saying the same thing in different words. For the heating of the head is nothing more than the warmth produced in it . . . and the wound nothing more than the cut to the flesh. (Galen, *On Containing Causes* 9. 6)

It is no good imagining that you have explained something merely by redescribing it. Molière explanations (see III.1*b* n. 8) are no explanations at all; and Aristotelian formal causes, like the early Stoic containing causes, perhaps court such vacuousness. On the other hand, while Galen is perfectly happy to ascribe the perfection of structure of the natural world to the excellence of the Demiurge's conception and forethought, he perhaps saw no need to elevate such things to the level of genuine, Platonic, hypostasized Forms, mysterious entities which in any case are liable to compromise the essential materialism of Galen's picture of the physical world.

4. Galen's Teleology

(a) Directed Teleology

Galen's acceptance of the modified Aristotelian causal taxonomy is evidence of a willingness to countenance teleological explanations in nature; and in fact he was committed to a particularly thoroughgoing teleology. *On the Function of the Parts* is devoted to demonstrating that every discoverable biological structure has an intended function, and that Nature and the Demiurge have arranged everything for the best. The most extreme

statement of this view occurs in a passage that Galen himself describes as a 'Hymn to Nature' (*On the Function of the Parts* 3. 236–44); but he regularly attacks other theorists (principally Erasistratus, Asclepiades, and their followers: §§2*a*, 3*b*, above) for calling Nature 'an aimless workman' (3. 364), and daring to hold that some organs might exist to no purpose.

Galen was clearly both familiar with and heavily influenced by the teleology of Aristotle's *Parts of Animals*.[38] But in his insistence that the world must be the result of the conscious efforts of an intelligent artificer, Galen leans towards the Plato of the *Timaeus* (III.4). *On the Function of the Parts* abounds with references to a Platonic Demiurge; and while Galen refuses to commit himself on the issue of the source of the arranging power in animals' bodies, holding this to be a matter for philosophers rather than doctors (*On the Formation of the Foetus* 4. 687–8), even in this passage Galen treats Nature personified as a proper object for worship.[39]

It is still argued by some that, in spite of the omnipresence of the Demiurge, Galen's teleology is Aristotelian in cast. Siegel (1973: 17) points to Galen's approval of Hippocrates' conception of nature's constructive power (*On the Doctrines of Plato and Hippocrates* 5. 791) as something which 'clearly reflects Galen's impersonal teleological interpretation of nature'. But this endorsement comes on the heels of Galen arguing (in imitation of a Stoic trope: VII.3*a*) that, just as it would be absurd, if you see a house or a ship but not the craftsman responsible for it, to conclude that it was spontaneously generated, so too when confronted by anything else clearly evincing design you must so attribute it even when ignorant of the artisan's nature and identity (5. 798–90).[40] Galen's Demiurge resists such easy demythologizing.

(b) The Scope of Teleological Explanation

Galen is well aware that his teleology is not Aristotle's (much less Theophrastus': V.3*b*). At the beginning of *On the Function of the Parts*, he tells us he felt impelled to discuss these issues since Plato and Aristotle had not fully understood the perfection of Nature's design (*On the Function of the Parts* 3. 16–21). In this vein, he praises the excellence of design of the gall-bladder (3. 372–4) which Aristotle thought to be a functionless residue (*Parts of Animals* 4. 2. 677^{a}12–19; cf. IV.3*e*; and compare his attitude to
end p.386

Erasistratus on the spleen: §2*a* above). Nature, for Galen, does nothing in vain in a variety of different senses. It will not produce a structure that exists throughout a species if that structure has no function to fulfil; but equally it will arrange the various structures thus created in the best (i.e. most economical) possible way.[41]

It is instructive to compare a particular case of Galen's teleology against its Aristotelian counterpart. Galen holds, like Aristotle (IV.3*a*), that facial hair is produced by the expulsion of the 'thicker residues' (*On the Function of the Parts* 3. 899–914), and he echoes both Aristotle and Plato in supposing that facial hair demands to be explained in terms of its protective value. But he is also committed, here as elsewhere, to the fundamentally un-Aristotelian view that these facts of human hirsute adaptiveness were consciously designed by a benevolent Creator (cf. VII.3*a*).

Indeed, he holds that a system's beauty can be part of its teleological explanation. Aristotle will occasionally invoke the notion of *to kalon* in connection with the explanation of nature (*Parts of Animals* 1. 5. 645^{a}23 ff.: IV.3*e*); and *to kalon* is sometimes rendered as 'beauty'.[42] But Aristotle's concerns here are not aesthetic ones, since he goes on to explain that it is because organisms have purposes that makes them (and indeed the study of them) fine and noble. Galen, on the other hand, frequently seeks to account at least partially for the existence of some structure as being an adornment, a move far more appropriate from within the context of a directed teleology:

410 moreover, the hair of the beard not only covers the cheeks, but also contributes to their ornamentation; for man seems more august, particularly as he grows older, if he is well endowed everywhere with hair.[43] For this reason too Nature has left the cheekbones and the nose smooth and free of hair; otherwise the whole face would be savage and bestial, and totally inappropriate for a civilized, social animal. . . . On the other hand, for women the rest of whose bodies are soft and hairless in the manner of children, the bareness of the face would not be unfitting; moreover, this animal does not have an august character as the male does, and so has no need of an august form. For I have already shown many times . . . that Nature makes for the body a form appropriate to the character of the soul. And the female kind needs no special covering as protection against the cold, since they generally stay indoors; they do, however, need long hair on their heads both for protection and ornament. . . . However, another function renders the hair on our chins and heads necessary for us. For since the exhalation from the humours rises to the head, Nature makes use of its thicker residues in particular to nourish the hair; and since

men have as much more of these residues as they are warmer than women, she has devised two ways for men to evacuate them. (Galen, *On the Function of the Parts* 3. 899–901)

Apart from its use of the concept of ornamentation, this passage is notable for its invocation of the appropriateness of body to soul. Elsewhere (3. 79–82, 208–9, 264, 848) Galen remarks that apes have ridiculous bodies to match their absurd souls. Yet he also holds that apes' bodies are particularly well suited to clambering rapidly up and down trees, which is their function (3. 209–10). These claims are not, first impressions notwithstanding, in conflict.[44] Relative to their functions and capacities, apes are well-designed; but from an absolute standpoint, their souls (and likewise their bodies) are absurd (see further §5*b* below).[45]

Equally (in Galen's view at any rate), women are generally inferior to men; yet they may still be excellent *qua* women. Some member of a natural kind *K* can be a good *K*, and yet need not be good in any more general sense. Even if no *K* is good (in the unrestricted sense), it may yet be good that there are *K*s. Galen's account here is, then, strongly analogous to the one which I sketched on the Stoics' behalf (VII.3*a*).

(c) The Limits of Creative Power

Galen's Demiurge is constrained (like Plato's: III.4*a*) by the intrinsic limitations of the material he has to work with. Galen (following Aristotle) thinks that eyelashes function as a protective palisade: yet they must be able to discharge this purpose without impeding other necessary functions (such as that of vision) in the process. Eyelashes are derivative in function, their usefulness depending on the prior functional utility of the eyes: they serve to protect the eyes, and would thus be pointless if the eyes themselves were functionless, or if they interfered with that function. But if the lashes are too thin they will fail to keep foreign bodies out; on other hand, if they are too thick or long, they will impede vision:

411 has our Creator then ordered these hairs alone always to maintain the same length, and do they maintain it as they have been enjoined either because they fear their Lord's command, or revere the God who commands it, or think it better to do so themselves? Is this how Moses reasons about Nature[46] (and it is a better way than Epicurus')? (Galen, *On the Function of the Parts* 3. 904–5)

end p.388

What, then, are the mechanisms (the *sunaitia*, as Plato would call them: III.4*b*, **144**) with which the Demiurge achieves his goal? Must we attribute intelligence and intentionality to the materials themselves, in order to avoid a universe of pure Epicurean chance which is manifestly incapable of throwing up anything of the complexity of the biosphere (ibid. 3. 74–9, 868–76)?

If those really were the only available possibilities, the random Epicurean world, for all its attendant difficulties, might well appear less implausible than the alternative of a directed teleology committed to that sort of panpsychism. However,

412 it is best for us to do neither, but, while maintaining the principle of generation from the Creator in all things generable, as Moses does, to add to this the material principle. . . . Since he has decided that it was necessary to make them [sc. eyelashes and eyebrows] so he placed under some of them a hard cartilaginous body and under the others a tough skin attached to cartilage by means of the brows. Now it was not enough merely to will that they were so: for he would be incapable of transforming a rock into a man all of a sudden even if he wanted to. This is where my teaching and Plato's . . . differs from that of Moses. . . . Moses believed everything to be possible for God, even if he should wish to make a horse out of beef and ashes. We, however, do not think that this is true, saying rather that some things, being naturally impossible, God does not attempt at all, but rather chooses the best course of action from among the possible. (3. 905–7)[47]

Galen's Platonic *via media* between random mechanism and universal creation *ex nihilo* allows him to generate an account of divine power strong enough to do the requisite explanatory work in resisting the explanatory lacunae of both atomism and Aristotelian immanent teleology, yet not one which is so strong as to fall foul of the more obvious Problem of Evil arguments (compare the Stoics: VII.3*a*; Plotinus, XII.2*d*).

The Creator wills everything to be for the best, thus demonstrating his perfect wisdom, power, and goodness (ibid. 3. 237–8). But even so, from the arrangement of the heavens to the architecture of the human body, the Creator's power is materially limited:

413 consider the matter from which each part is constructed, and do not imagine that from menses and semen[48] an immortal object could be produced . . . as bright and fair as the sun. (3. 238)

moreover,

414 are you willing to have the sun formed from the substance of blood, so prone to putrefy and so filthy? . . . I will not, indeed, show you the sun in the

end p.389

body of an animal, but I will show you the eye, a very brilliant instrument, resembling the sun as closely as is possible in the body of an animal. (3. 242–3)[49]

The opponents of teleological explanation who object that the world is not as good as it might be demand a causal impossibility; and as soon as one accepts that the Creator really is subject to material restrictions, the claim that we inhabit the best of all causally possible worlds becomes a good deal easier to swallow. A little earlier, Galen has castigated the anti-teleologists for failing to see that true excellence is formal, not material, in nature: a sculptor's clay maquette exhibits as much beauty of form as the realized marble or bronze, and the Demiurge's skill is just as evident to the discerning eye even in such inferior material (3. 237–40).

(d) The Regulative Function of Teleology

Thus Galen rejects the mechanism of the Atomists as explanatorily inadequate, as indeed it is unless and until it can be supplied with a detailed account of the sort of cybernetic, feedback processes that can allow mechanism to mimic intentionality. But Aristotelian teleology, where final causes are irreducible parts of the force of nature itself, yet a nature which involves no consciously entertained end, appears to be at least equally inadequate. For what could such a force amount to? All the obvious examples of intrinsic purpose invoked by the Peripatetics involve consciously entertained goals. At *On Antecedent Causes* 6. 56–60, Galen gives a series of examples of final causes along Aristotelian lines: carpenters craft beds and weavers weave coats because such goods are in demand, that is they are desired. If nature too involves purposes, it should exhibit a similar structure.

Moreover, Galen regularly attacks other doctors (notably Erasistratus and Asclepiades: §§2*a*, 3*b* above) for failing to appreciate the extent of nature's teleology. Animals' bodies can contain no parts which are super- fluous, even if such parts may, in the Aristotelian way, be accounted for as necessary residual outcomes of further processes which are in themselves useful, since a benevolent and skilful Demiurge would surely have put such residues to some use, even if only an ornamental one.[50] Erasistratus and Asclepiades even dare to hold that certain organs are simply useless and cannot even be explained as being the inevitable by-products of other, necessary, functions (§2*a* above). Galen treats such views with immense

and characteristic scorn. Peripatetics and Erasistrateans share nothing other than the principle that nature does nothing in vain; yet

415 even as regards this doctrine, their agreement is merely verbal. In practice, Erasistratus destroys it on countless occasions. For, according to him, the spleen was made in vain, the omentum in vain, the arteries that originate in the kidneys in vain. (Galen, *On the Natural Faculties* 2. 91[51])

Galen charges Erasistratus' ignorance of anatomy with the blame for this shortcoming (anatomy is not entirely, or even principally, a practical enterprise for Galen: see further §5*b* below). Detailed anatomical research serves to confirm the view that nature is indeed providential in its arrangements by emphasizing their complexity and well-adaptedness. Galen effectively accepts the Platonic argument from design ([AD]: III.4*a*):[52] the natural world must be the product of conscious purpose, executed by a superior intelligence, since no other general account of their systematic nature is plausible; but such a superior, benevolent, and powerful intelligence will not simply leave things lying around to no purpose if something better can be done with them.

Galen's Creator is not, as we have seen, literally omnipotent; nor indeed (and here Galen goes further than Plato) is he even omnicompetent. Galen explicitly allows (*On the Function of the Parts* 4. 351–62) that the Demiurge will sometimes, like any craftsman, blunder: he discusses the case of individuals born with six fingers on one hand (even though, as he explains at great length, five is the ideal number of fingers). But the mistakes of such a superior intelligence will be rare and relatively insignificant; and they will not be the general, uniform errors in the construction of entire species which would undermine the plausibility of the directed teleology in the first place (4. 355).[53]

But if the best overall explanation is one involving directed teleology, Galen must be able to demonstrate that, first appearances notwithstanding, even apparently useless organs fulfil some function. Thus Galen's teleological commitments amount to a methodological rule: all apparently functionless organs must be thoroughly investigated until they reveal their hidden purposes. Furthermore, an intelligent Demiurge will choose the best and most economical engineering solution to any particular problem. If possible, functions will be combined and the amount of material expended reduced; parsimony is a creative virtue, because among other things it is an artistic and technical virtue.[54]

end p.391

5. Powers, Functions, and Activities

(a) Function and Anatomy

And, Galen claims, anatomical research lends this position powerful empirical support. The works of Nature are, when properly analysed, enormously more complex than first appearances suggest. Consider the insertions of tendons into the human finger-joints (*On the Function of the Parts* 3. 77–8). There are, Galen says, thirty joints in the ten fingers, three in each finger; and since each joint generally boasts four tendon-insertions, one at the front, one at the back, and one on each side, one might suppose that the total number of such insertions was 120. But in fact there are only 118, since the first thumb-joint lacks the interior insertion. This is no accident, given the necessary functions the hand must perform:

416 certainly if we flexed this joint in the same way as the others, I know you would then harshly and vehemently criticize the uselessness of Nature's labour in creating an unserviceable motion, and a tendon that was superflu- ous. (ibid. 3. 77)

An excellent Nature must supply all and only those functions which any particular animal requires. Thus the functional well-adaptedness (or rather createdness) of the hand is explicable in terms of a directed teleology (and perhaps only thus), while supplying practical evidence for the hypothesis (compare Chrysippus on divination and determinism: VII.2*d*).

The science of biology progresses by means of a concentrated investigation into the relations between form and function, on the assumption that nothing will have been done by nature in vain. At the beginning of his great treatise on anatomy, Galen writes:

417 the study of anatomy has one use for the scientist who desires knowledge for its own sake, another for him who desires it not in itself but in order to show that nothing comes to be at random from nature, another for him who arms himself from anatomy with premisses relating to the knowledge of some activity (*energeia*), be it physical or mental, and yet another in addition to these for the person who has to remove splinters and arrowheads well, or to cut something out in the appropriate way to deal surgically with ulcers, fistulae, and abscesses properly. (Galen, *On Anatomical Procedures* 2. 286)[55]

end p.392

Those ignorant of practical anatomy will be more inclined, with Erasistratus, Asclepiades, and their followers, to suppose that some structures exist to no purpose; but detailed anatomical investigation will serve to disabuse them of that. At *On the Function of the Parts* 4. 348–50, Galen describes his investigation into the elephant's trunk: at first you might simply suppose it to be a useless appendage, until you see how it is employed in lieu of a hand; moreover, it is also a nose, and its economical internal design may be revealed by dissection.

But mere dissection is not enough: one of the uses of anatomy listed in **417** involves the discovery of *energeiai*, the characteristic activities of the various parts and structures in the human body (impediment to which is the form of disease: *On the Therapeutic Method* 10. 50–1, 59–61, 63–7, etc.). And crucially, these cannot invariably be inferred simply from the morphology and relative location of the structures themselves. Although, Galen thinks, it is a reasonable inference that anything which is duct-shaped must be designed for carrying something (*On Anatomical Procedures* 2. 727), inferences of that sort are in themselves quite incapable of supplying a complete and detailed knowledge of functions. Rather

418 that part of anatomy dealing with dead animals teaches the position, number, and particular substance of each part, as well as its size, shape, and composition; that of the living sometimes teaches the activity immediately, and sometimes the premisses for deducing it. (ibid. 2. 707–8)

In other words, vivisection is required. But, whatever the case for Herophilus and Erasistratus (IX.1*b*–*c*), vivisection of human beings was not an available option for Galen. As a result, he was forced to resort to the vivisection of animals; and since, as he frequently emphasizes, his primary goal is not the discovery of function and structure in animals at large but in human beings, he needs a theory which will allow him to infer to the hidden operations of the human body on the basis of what he discovers in animals. To this end he takes over and refines the notions of structural homology and analogy which derive ultimately from Empedocles (who compared hair with feathers and scales: 31 B 82 DK; cf. 39 B 79, 83 DK), by way of Aristotle.

He distinguishes rigorously between parts which are genuinely homologous across species, differing only in minor and superficial manners from one animal kind to another, such as the heart:

419 all doctors and philosophers who are skilled in natural science agree that the activities are consequent upon the particularities of the substances. Every part, then, which is substantially the same acts in the same way, even in animals that differ widely. . . . Each heart, then, has the same activity as every other, as does each thorax and each lung. (Galen, *On Anatomical Procedures* 2. 611)

end p.393

Genuinely homologous structures are that way since they have been created to fulfil precisely the same role in each animal, morphological variations notwithstanding:

420 if the organ in question is an instrument of vision, it is an eye even if it is differently constituted, for example in men and crabs; and if it functions as an instrument for walking, it is a leg whether it belongs to an elephant, a sheep, a goat, or a man (Galen, *On the Doctrines of Hippocrates and Plato* 5. 203)

It is thus only when different species have distinct functions to fulfil that we may expect their organs to be substantially different, even though they may yet exhibit similarities of analogy (as feathers do to hair and scales). On the basis of such general considerations (as well as empirical evidence) Galen develops a classification of the 'six kinds of animal which are closest to man'.

The six kinds are mentioned in several places (mostly in *On Anatomical Procedures*: 2. 375, 422–3, 430, 547–9, etc.), and, some obscurities aside, it is clear that Galen intends to set up a taxonomy of animal parts (rather than species as such, since the relative ordering of the six kinds varies from structure to structure, although apes are invariably the first group[56]) in order better to be able to infer to the hidden structures and operations of the human body. In the case of the vital organs, such as the heart and lungs, there will be no significant differences; equally, since voice production in itself is specifically the same in all animals that vocalize, what can be learned about the larynx from experiments on pigs will transfer without remainder to the human being. Other cases, where the family relationship is more distant, will require analysis of the sequence of transitional forms[57] in animals as they more closely approach the human in general anatomy and function, ending up with rhesus monkeys and Barbary apes, which mimic the human form most closely (cf. §4*b* above). Thus armed with a proper understanding of the family relationships of nature, we may infer from the results of animal experiments to conclusions about the specifically human condition.

Galen leaves us detailed records of his experiments on live animals throughout *On Anatomical Procedures*: the most famous involve determining the effects on the voice and the motor functions of cutting and ligating various nerves around the spine of pigs (2. 661–706 Kühn; 11. 10–11, 127–34 Simon; 14. 7–8, 262–73 Simon;[58] cf. *On the Doctrines of*

end p.394

Hippocrates and Plato 5. 232–5), as well as investigations into the brain and the spinal column (*On Anatomical Procedures* 9. 10–13, 13–31 Simon). In some cases it is only by actively intervening to interfere with some activity's function that one can properly determine what it is. In *On the Doctrines of Hippocrates and Plato*, at the end of his extensive and brilliant refutation of the Stoic doctrine that the heart is the source of impulse and control in the body,[59] Galen writes that, in order to determine whether the heart transmits power to the brain,

> **421** one must investigate by experiment on animals what sorts of structure connect the heart with the brain and how many they are, and then cut or crush or stop them with ligatures at the neck, and observe the effects on the animal. (Galen, *On the Doctrines of Hippocrates and Plato* 5. 263)

The results of such detailed investigations will reveal that there are only three such structures (veins, arteries, and nerves), and that interfering with the first two has no effect on functions of any kind, while ligating the nerves between heart and brain affects neither the pulse (hence the brain is not transmitting power to the heart), nor general motor control (hence the heart is not transmitting power to the brain): 5. 263–5. The nerves transmit the power involved in motor-control; they have their origin in the brain (since they ramify out from it); moreover, experiments with section and ligation can show the direction of transmission of that causal power, since all control is lost downstream of the break (cf. 5. 519–21; *On the Movement of Muscles* 4. 371–3). Throughout his demonstrations, Galen supposes that the causal powers in question must be transmitted either directly by contact or through an intermediary vessel; and careful anatomical inquiry will serve to reveal the nature and organization of those vessels in the living animal.

(b) Faculties and the Soul

Galen is not, for all that, unwilling to postulate the existence of certain causal 'powers' or 'faculties' (*dunameis*; cf. Diocles and Herophilus, IX.1*a–b*) that belong to particular organs and whose causal operation is more mysterious. Indeed, he wrote a whole text (*On the Natural Faculties*) defending the need to introduce such hypotheses to account for the fluid dynamics of the body against the pure mechanists; and the various metabolic processes require explaining in terms of more than merely the types of mechanical principle employed by the Atomists and those influenced by them.
In particular, Erasistratus' reliance on the principle of *horror vacui* (IX.1*c*; cf. V.2*b*) is mercilessly attacked (*On the Natural Faculties* 2. 63,

98–9), and the inadequacies of the atomist account of magnetism in terms of the physical interactions of minute particles held up to withering ridicule (ibid. 2. 44–53; cf. Lucretius 6. 906–1089). Galen holds it necessary to posit four basic 'faculties' for various parts of the body—attraction, retention, propulsion, and expulsion (*On the Natural Faculties* 2.

6–7). Only thus can we hope to account for the differentiation that various parts of the body are capable of making among the various things they attract (the predilection of the bladder for urine, for example, or of the stomach for food). The various mechanical principles adduced by the opposition are simply not fine-grained enough to account for such selectivity; moreover, it is impossible on atomist principles to account for the specific epispastic properties of certain drugs, antivenins, and salves (ibid. 2. 40–4, 53–5, 161–2).

Galen apparently here indulges in precisely the sort of Molière explanation which he elsewhere takes pains to avoid (§3*d*, **409** above). What are these so-called faculties except mere redescriptive nominalizations of the effects in question? Galen is alive to this charge:

422 (i) all the . . . faculties fall under the category of relations: they are primarily the cause of the activities [sc. of the various organs] but incidentally [cf. §3*c*, **404** above] of their effects. (ii) But given that the cause is relative to something, being the cause *of* what results from it alone, and of nothing else, it is obvious that the faculty also falls under the category of relation. (iii) And so long as we are ignorant of the essence of the activating cause we call it a faculty: thus we say that there is in the veins a blood-producing faculty, a digestive faculty in the stomach, a pulse-creating faculty in the heart, and in each of the other parts a specific faculty corresponding to its activity. (iv) So if we are to investigate methodically how many and what sort of faculty there are, we need to start from the effects—for each effect comes from a specific activity, and each activity from a specific cause. (Galen, *On the Natural Faculties* 2. 9–10)

For Galen, then, hypothesizing some faculty or other is merely the beginning of science, and not its end, as it is in Molière explanation. Faculties are place-holders, inferred entities postulated on the basis of given effects. But the investigator must proceed to determine what precisely the faculty consists in, which in turn consists in pairing it with an activity (*energeia*) of the animal in question. These *energeiai* are themselves inferable from their observable effects: **422**(iv) draws the methodological conclusion that investigation from such effects to isolable causes is a feasible procedure.

Sometimes it will turn out that we are not really in a position to say much about the real nature of the function, other than that it exists and operates in a certain sort of way. Galen offers no account of *how* the bladder exerts its urine-attracting properties. And in general he is loath to commit himself on questions whose solution is, for whatever reason, empirically unavailable:

423 that the majority of disagreements in philosophy have not been concluded is not surprising, since these issues are not susceptible of clear judgement by empirical test (*peira*); for this reason some assert the universe to be ungenerated, others that it had a beginning, just as some say that there is nothing outside it enclosing it while others hold that there is, some of the latter claiming that it is void and contains no substance at all, while others say that there are other universes uncountable in number, a multitude stretching to infinity. It is impossible to adjudicate such a dispute with clear perception. (Galen, *On the Doctrines of Hippocrates and Plato* 5. 766)

Elsewhere, Galen frequently speaks of the importance of subjecting theoretical claims to empirical testing: indeed it is the core of his synthesis of empirical and rational medicine (cf. IX.2); reason suggests the therapies, which must then face the tribunal of experience (cf. e.g. *On the Therapeutic Method* 10. 29, 393–4, 405–6; *On Hippocrates' 'Nature of Man'* 15. 152–3[60]). But there are some questions which simply cannot be settled in this way, and must be left to speculation, such as the substance of the gods (although Galen does not include among such indeterminable cosmological questions that of whether the world is providentially organized by a divine artificer: *Doctrines of Hippocrates and Plato* 5. 780–1; cf. 783–91).

Equally resistant to discovery is the real nature of the soul:

424 that every body in our part of the universe [i.e. the sublunary realm] comes to be as a result of the mixture of elements I hold to be securely known . . . but whether as a result of the complete intermixture of the elemental bodies themselves, or whether of their qualities only, I think it unnecessary to know, and make no pronouncements on the matter (although I think it more plausible that the mixtures are mixtures of qualities). But as for the soul, whether it is immortal and directs animals in conjunction with bodily substances, and whether there is any substance of the souls as such, I assert that these things cannot be securely known. (Galen, *On the Substance of the Natural Faculties* 4. 762–3)[61]

But fortunately it does not matter for practical purposes whether the soul is substance or property, mortal or immortal, corporeal or incorporeal, and separable or inseparable from the body (ibid. 4. 763–4; *Doctrines of Hippocrates and Plato* 5. 794–5; *On the Substance of the Natural Faculties* 4. 759, 763–4). These questions have no bearing on the fact *That the Characteristics of the Soul Depend upon the Temperaments of the Body* (4. 767–822; see further § *c* below), even though the evident fact that physiological conditions have psychological effects makes it hard to see how it

end p.397

could fail to be corporeal (ibid. 4. 782, 787–8), while the notion of psychic immateriality is hard to make sense of (4. 776–7). The soul, then, whatever its real nature, is to be thought of in terms of its capacities and functions:

425 everyone knows that we possess souls, since we plainly see the things that are activated through the body: walking, running, wrestling, the many varieties of perception; and we know on the basis of an axiom that commends itself naturally to all of us that there is some cause for these activities: for we know that nothing occurs without a cause. But because of our ignorance as to exactly what the cause of these things might be, we assign it a name on the basis of its capacity to do what it does. (Galen, *On the Substance of the Natural Faculties* 4. 760)[62]

Thus Galen refuses to commit himself to the view that the soul actually *is* the psychic *pneuma* (*Doctrines of Hippocrates and Plato* 5. 605, 609), although he thinks that it is demonstrable that its power is communicated in the form of *pneuma* via the nervous system, the existence of the powers in question guaranteed by the non-negotiable causal principle GP1 (§2*b* above); and it is overwhelmingly likely that they will be mediated by discoverable physical connections.

Yet for all his commitment to a priori causal principles, Galen never loses sight of his fundamental empiricist commitments: and he will on occasion conscientiously note phenomena which are difficult to account for from the standpoint of his own theories. He thinks that certain types of causal influence (the paradigm case is that of the transmission of neural impulses to the extremities of the body) are too rapid to be accounted for on the assumption that material parts themselves carry the impulse:

426 just as those who have chosen to move their toes do so immediately, not because the part itself possesses reason, but because a certain power is transmitted to it instantaneously, so too it occurs in the case of the heart and the arteries. The rapidity of the motion is consistent with the faculties, but tells against material substances. (Galen, *On Whether the Arteries Naturally Contain Blood* 4. 735[63])

Thus, while he never abandons his basic humoral pathology, he will allow more exotic modalities of causal transmission than most of his predecessors. In his late treatise *On Affected Parts* Galen notes that apparently insignificant causal influences (in this case retained sperm and menstrual fluid) can have large-scale systemic effects, a view with Hippocratic antecedents (II.1*b*, **74**; cf. **63**, **73**), and one obviously compatible with his own account of the structure of causation (§2*b* above). He points out first that

end p.398

small amounts of poison can effect great changes, as in the case of the venom of spiders and scorpions (where sometimes the puncture in the skin is so small as to be virtually invisible: *On Affected Parts* 8. 421), and rabies (8. 423; cf. IX.3*b*); but he goes on to remark that

427 some even think that some substances can alter things that are close to them just by contact and solely by the power of the transmitted quality. This can easily be seen in the marine torpedo-fish,[64] which possesses a force so strong that by transmitting this alterative power through a fisherman's trident to his hand it can immediately induce total numbness in it. These are sufficient indications that a small substance can effect large alterations by contact alone; and no less is the case of the Heraclean[65] stone, which they call a magnet; for iron which is in contact with it is suspended from it without attachment, while a second piece in contact with the first piece that is in contact is also suspended in a similar manner to it, and even a third from the second. (Galen, *On Affected Parts* 8. 421–2)

We noted above that Galen rejects the Atomists' explanation of magnetic attraction as absurd; but equally he admits that he has no real account of how such attractions operate to offer himself. Yet they cannot simply be wished out of existence, in the manner of the armchair physicist and physician. The proper scientific course is to accept the phenomena and work, however tentatively, towards an account of them, admitting that one's explanations may be provisional only. The following passage is typical:

428 the heart itself, having all the faculties of attraction that one might think of, snatches and as it were gulps down the inflowing material, receiving it rapidly into the hollows of its chambers. For perhaps the heart possesses the same faculty (only to a greater extent) as the bronze-worker's bellows which draw in the air by dilation; or perhaps the heart (being as it is the origin of the innate heat) does not lack the faculty possessed by the flames of lamps which draw up the oil; or perhaps its acts as the Heraclean stone attracts iron, by appropriateness of quality: for what could be more appropriate for it than air for refrigeration? (Galen, *On the Function of the Parts* 3. 481)

Thus Galen's account of the nature of physical interaction (and, by association, the causation of disease) is fluid, and responsive to the results of empirical investigation. But it is informed throughout by Galen's fundamental belief in the necessity of postulating a benevolent, organizing creator of the world: if his substance is in doubt, his existence is not. Yet for all that, Galen is still strongly drawn to the idea that the ties that bind the world together must ultimately involve contact and be physical in nature.

end p.399

His hypothesized powers are a far cry from the immaterial causation that is to be the cement of the Neoplatonist universe (cf. Iamblichus, XII.3*b*, **471**).

(c) Action and Responsibility

As was noted above (§2*a*), Galen repudiated Chrysippus' account of the nature and structure of the emotions at length (*On the Doctrines of Hippocrates and Plato* v. 369–91); the emotions cannot simply be judgements writ large, and passions are not just errors. Indeed, Galen wrote two complementary pamphlets *On the Passions of the Soul* (v. 1–58) and *On the Errors of the Soul* (v. 59–103), rigorously distinguishing between them. In Platonizing fashion, he locates passion in a distinct psychic faculty in the heart from reason and judgement, which are to be found in the brain; and

429 errors arise from false belief, while passions are caused by an irrational power within us which refuses to obey reason. (Galen, *On the Passions of the Soul* 5. 2–3; cf. 7)

Galen is primarily interested in such phenomena, and their distinctions, from a therapeutic point of view: and the pamphlets on passions and errors are in part dedicated to the project of showing how we may become better people. The main point is that, having first recognized that we have faults that need eradicating (ibid. 5. 3–6), and with the assistance of someone we trust to help point out our faults to us (5. 8–14), we may, provided that we devote constant and unflagging efforts to the project, seek to eradicate, or at least domesticate, those aspects of ourselves that we have come to dislike (a proneness to anger, for example): 5. 14–21. Eventually, after long training, we come first to suppress the effects of the emotions, and finally, as a result, get rid of our disposition to excessive emotional reaction, which in turn removes one of the most potent causes of errors in judgement.[66]

Thus, Galen thinks, a cognitively based regimen can serve to produce beneficial changes in character; and he wants to hold us responsible for the way in which we choose to live:

430 man alone as compared with other things has the special gift of reason: if he casts this gift aside and indulges in anger, he is living and acting like a wild animal rather than a man. (5. 22–3)

But elsewhere Galen seems committed to the view that mental phenomena are at bottom physical; or if this is too strong, at the very least that many psychological dispositions have physiological causes. This is, indeed, the burden of his treatise *That the Characteristics of the Soul Depend upon the*

end p.400

Temperaments of the Body. And the passions are conceptualized, as they are so frequently in the ancient world, as alien forces which take control of individuals against their will. Control of the passions involves putting reason back on its throne by undergoing the type of psychotherapy outlined above: yet will the decision to undertake the psychotherapy itself, and the rigour with which one decides to carry it through, itself not be caused by, or at the very least conditioned by, physiological states? And if so, how can these actions be free?

Galen adopts the view, ultimately traceable to Plato's *Republic*, that true freedom consists precisely in reason's fulfilling its proper controlling role. An individual is free, and hence acts freely, just in case his reasoning powers are not subverted by emotion and desire; freedom is freedom *from* the passions. And that story is, of course, compatible with psychological determinism. For Galen, unlike Alexander (X.4*c*), freedom does not require that the free agent have genuinely open choices, in the sense of distinct options which are, as yet, causally undetermined. There is, then, for Galen no Aristotelian problem of how, if we are not somehow in control of how things appear to us, we may be said to be responsible for what we do (IV.4*b*, **177–8**). Moreover, differences of character, in virtue of which we will become either virtuous or vicious, are largely innate (*Characteristics of the Soul* 4. 768–9, 816–19; *On the Passions of the Soul* 5. 37–40): education and training can only achieve what the raw material it goes to work upon will allow. This has obvious implications (ones which Alexander at least would consider fatal: X.4*c*; cf. VII.2e, **280**) for our notions of responsibility and desert:

431 everyone comes to be what they are because of their bodily temperaments. How then, they say, can someone be justly praised or blamed, hated or loved, when they have become either wicked or virtuous not as a result of themselves but on account of their temperament, which is clearly the result of other causes? Because, we reply, it belongs to all of us to rejoice in the good and to pursue and love it, and to turn away from the bad and detest it and flee from it, paying no regard to whether it is generated or not. Thus we destroy scorpions, poisonous spiders, and vipers even though they have become what they are by nature and not as a result of themselves. Consequently, we can with justification hate the wicked among men without further calculating that they are the active cause of their becoming thus. Furthermore, we put the irremediably wicked to death for three good reasons: so that they cannot wrong us while they are alive; so that those like them will be afraid lest they too be punished by those they have wronged; and because it is better for these people themselves to die, given that their souls are in such a state of corruption that not even the Muses themselves could educate them, nor could they be improved by Socrates or Pythagoras. (Galen, *That the Characteristics of the Soul Depend upon the Temperaments of the Body* 4. 814–16)

end p.401

Thus Galen simply denies that the rationale for punishment (and hence also for praise and blame) depends upon establishing responsibility and hence desert in any strong sense. The standard justifications for punishment (public safety, deterrence, and elsewhere at least rehabilitation) work perfectly well regardless of whether the world is deterministic or otherwise; and the same is also true of retribution. We destroy the irremediably reprobate simply because they are so, and because their lives are objectively speaking worthless. There are, of course, difficulties with that latter, harsh view: but at the very least Galen shows that, *pace* the Epicureans and Aristotelians like Alexander, there is nothing incoherent about the notions of praise and blame, honour and punishment, in a deterministic world.

6. Conclusions

Ptolemy developed the most sophisticated geocentric ancient astronomical model, and tried to show how it could be given a physical realization (§1*a*). He also argued for a relatively sophisticated astrology, one immune to the more obvious sceptical counter-arguments (§1*a*; cf. VIII.5).

His near-contemporary Galen elaborated a comprehensive account of causation. In refutation of Erasistratus (§2*b*; cf. IX.2*c*), he urged that antecedent causes, even if not on their own sufficient for their effects, were none the less causes, since in concert with suitable physical dispositions they do condition their effects (§2*b*); these dispositions are preceding causes, which when activated by the external antecedent conditions produce the immediate necessary and sufficient containing causes of diseases (§2*c*).

Galen also makes room, in the syncretizing fashion of his age, for Aristotle's four causes (§3*a*), but with the addition of the Middle Platonist instrumental cause (§3*b*), and the general omission of the formal cause (§3*d*); and in pursuit of a rigorous classification (and the avoidance of fallacy), he distinguishes incidental from essential causes (§3*c*).

He further adopts a generous directed teleology from the Platonic tradition, complete with an artisan-God (§4*a*), although his application of teleological principles to physiological explanation owes more to Aristotle (§4*b*); but while he goes beyond what Aristotle would have accepted, Galen's artisan-God is still constrained by the ties of material necessity (§4*c*). Teleology also provides a methodological incentive to investigate the functions and powers of things (§§4*d*, 5*a*–*b*), powers which are empirically determinable, but must be accounted for within the structure of theory.

end p.402

Finally, Galen contributes to the debate on freedom and responsibility by suggesting that, even if our actions are determined by causes outside our control, we may none the less justly be held responsible for what we do simply because we are the sort of people that we are, regardless of the provenance of our characters (§5*c*).

end p.403

XII The Neoplatonists

R. J. Hankinson

1. The Roots of Neoplatonism

(a) Plotinus: Life and Sources

Plotinus (ad 205–70) is a mysterious figure. His pupil Porphyry (*c.*232–302) wrote a *Life* of his master: yet Porphyry himself confesses (*Life of Plotinus* 1) that he knows nothing of Plotinus' family or ethnic background. Plotinus took up philosophy at the age of 28 in Alexandria under Ammonius Saccas (ibid. 3).[1] In order to acquaint himself with the practices of the Eastern mystics, he took part in a disastrous military expedition against the Persians in 242–4, after extricating himself from which he settled in Rome, where he founded a philosophical salon in the house of one Gemina (9). Porphyry arranged his fifty-four extant treatises, not without certain tendentiousness, into six 'Enneads' (or groups of nine), in an attempt to exhibit the structure and coherence of the Plotinian system (24–6).

> **432** Those who have given evidence of their seriousness in writing by the great number of problems they tackled and their employment of an original manner of inquiry are Plotinus and Gentilianus Amelius. The former, I believe, has given an exposition of Pythagorean and Platonic principles with more clarity than any of his predecessors; the writings of Numenius, Cronius, Moderatus, and Thrasyllus do not even approach the precision of his writings on the same issues. (20. 69–76)

This was the view of the critic Longinus. Yet Plotinus was accused of plagiarizing his system from others, notably Numenius (21), a charge Porphyry is anxious to refute. Amelius wrote a book demonstrating Plotinus' originality (17), and Porphyry preserves a letter of his maintaining that these allegations (as well as those of unintelligibility and making his first principles of the worst kind) derive from the accusers' own glib incomprehension and a desire to mock (17–18).

Longinus speaks in **432** of Plotinus as expounding Pythagorean as well as Platonic doctrines; and although by this period the histories of Platonism and Pythagoreanism have become inextricably intertwined, we should

end p.404

glance briefly at the neo-Pythagorean doctrines of the period in order to determine the extent to which Plotinus was at least influenced by them. The study of earlier writers was an integral part of the intellectual life of Plotinus' circle:

433 at the meeting the commentaries of Severus, Cronius, Numenius, Gaius, and Atticus,[2] as well as the Peripatetic writings of Aspasius, Alexander, and Adrastus, and others that were to hand, were read out loud. (14. 101–4)

Porphyry indeed allows that there is much of Stoicism and Aristotelianism to be found in him. But he emphasizes that Plotinus did not merely read and comment in a dry scholastic manner on these works: rather he bent them to his own original ends.

(b) The Neo-Pythagorean Background

The most important figure of the Pythagorean revival is Numenius, who probably flourished in the middle of the second century ad .[3] Substantial fragments survive of his lengthy treatise *On the Good*, in which he insists on the incorporeality of Being (*to on*: Fr. 6 dP;[4] cf. Fr. 7 dP; Being and the Good are interchangeable for him: Frs. 16 and 19 dP), as well as its unchangeability and eternality (Frs. 7–8 dP). Most striking is his postulation of a tripartite sequence of gods:

434 the person who wishes to have understanding of the First and Second Gods must first of all distinguish everything into its appropriate position and ordering. . . . The First God, being in his own place, is simple and could never be divisible because he consorts entirely with himself. The Second and Third Gods are, however, one, but by being in contact with matter (which is the Dyad), he unifies it while being divided by it, which has a fluid and desiderative character. So not being in contact with the intelligible (for then he would be in contact with himself), and because he looks towards matter and is concerned with it, he becomes careless of himself. He comes into contact with and ministers to the perceptible, and in his desire for matter draws it up towards his own condition. (Numenius, in Eusebius, *Preparation for the Gospel* 11. 18. 1–5 Fr. 11 dP)

Thus the Second God, who exists and operates below the perfect self-contemplation of the good and unified First God, is drawn down towards the dubious realm of the sensible world by desire, in a kind of fall from grace. The First God is described as the Father of the Creator God
end p.405

(Fr. 12 dP), a fragment which emphasizes his complete detachment from the world. The Second, or Creator, God occupies an intermediate position between the higher life of the intellect and the more banausic involvement with matter. Numenius tries to clarify the relations between them:

435 everything which is given and comes from a giver passes to a receiver (slaves, riches, money), at least in the case of mortal and human gifts; on the other hand are such that, while they are transmitted from there to here, and although they arrive here, they do not leave there, benefiting the one without harming the other. . . . This noble gift is noble understanding which benefits him who receives without leaving the giver. It is the same as when you see a lamp lit and, deriving its light from another lamp, it does not take away the other's light, but simply has its own matter ignited by the other's fire. (ibid. 11. 18. 15–16 Fr. 14 dP)

The teacher imparts knowledge to the learner without thereby suffering any diminution of his own knowledge; and this provides a model for the way in which power is induced in lower beings by those higher in the scale, without in any way affecting the latter. This notion, which is to become central to Neoplatonism, amounts to a rejection of conservation principles in regard to causation (at least of this type; contrast Galen's GP6: XI.2*b*); and the analogy of the induction of knowledge is taken over by Plotinus himself (*Enneads* 4. 9. 5. 4–9; cf. 3. 9. 2).

But the actual structure of Numenius' hierarchy is still obscure. Proclus (*On the Timaeus* 1. 303 Fr. 21 dP) reports that Numenius parcels out the role of creation between the First and Second Gods, and accuses him of thereby confounding the primary reality, the Good, with inferior causes; but this interpretation seems mistaken.[5] Another fragment clarifies matters slightly:

436 if essence (*ousia*) and Form are intelligible, then Mind (*nous*) will be agreed to be prior to and the cause of it, and will be found to be itself the only Good. If the Creator God is the principle of becoming, then it suffices for the Good to be the principle of being; and just as the Creator stands in relation to it and imitates it, so too does becoming stand to being. And if the Creator of becoming is good, then surely the Creator of being will be the good-in-itself, this being part of its essence. For the Second, being double, creates of itself both its own form and the world, being a Creator, since the First is wholly contemplative.[6] So we may conclude by establishing four names for the four

end p.406

things: the First God who is good-in-himself; the Creator, who is good in imitation of him; the essence, which is one thing for the First, and something else for the Second; in imitation of which is the beautiful world, made beautiful by participation in the beautiful. (Numenius, in Eusebius, *Preparation for the Gospel* 11. 22. 3–5 Fr. 16 dP)

The highest god is equated both with the Good and with Mind (in contrast to Plotinus' system: §2*a* below; cf. Fr. 17 dP). The Creator is good only by participation in the First God, who is wholly self-sufficient, and at rest, although that rest itself takes the form of a continuous internal motion which is responsible, albeit at several removes, for the world's stability (Fr. 15 dP; cf. Heraclitus, I.3*b*, **30**; and the Stoics, VII.1*b*). Thus Numenius' Creator is only derivatively, and hence less than wholly, good (Frs. 19–20 dP), which accords with his being driven by his desire for matter (**434**). The 'four things' of the last sentence are puzzling, given Numenius' insistence elsewhere upon a trinity. The odd one out is the essence, which seems to be a bridging device between the First and Second Gods rather than an individual substance, [7] while the Third God (the world) is the material aspect of the Janus-faced Second God.

Numenius' metaphysical fundamental concern is to keep open the lines of communication between his levels of reality, presumably since it would otherwise be impossible to explain how the higher forms could have any effect at all at the lower levels. Thus Fr. 19 maintains that lower beings participate in the highest by thought alone, while thought itself is the sole possession of the First God. However,

437 Numenius equates the First with that which is alive, and says that it thinks with the additional aid (*proschrēsis*) of the Second, while the Second is assimilated to Mind, and creates with the additional aid of the Third, which is assimilated to practical thought. (Proclus, *On the 'Timaeus'* 3. 103 Fr. 22 dP)

This report, admittedly from the unfriendly Proclus, seems to have things the other way around; 'practical thought' (*dianoia*) contrasts with pure abstract intellectual activity, which seems to be primarily the preserve of the Second in the hierarchy. But whatever we are to make of that, [8] the elements in Numenius' scheme are tightly interrelated. The First God can only think with the aid of the second, to whom thought is proper. As **436** stresses, their essences are distinct; equally, the practical thought required in order to create a cosmos is something that the Second can only achieve by calling into play the Third God with its desire for matter (**434**), practical thought being, as Plotinus will say, 'the job of Soul, not of Mind' (*Enneads*

end p.407

3. 9. 1, 35). However, this mutual interdependence of the principles is thoroughly un-Plotinian: in particular, Plotinus consistently (e.g. 3. 8. 11) rejects the idea that the higher, being self-sufficient, might in some way require the lower.

A long passage of Calcidius'*On the 'Timaeus'* (295–9 Fr. 52 dP) preserves Numenius' 'Pythagorean' doctrine of matter, contrasting it with both Stoic and Platonist accounts. Matter, at least in its original undifferentiated state, is equivalent to the Indefinite Dyad (cf. I.4*a*; III.4*d*; X.1*a–b*, **347**, **352**), and is contrasted with the Monad, or God, who makes it determinate (295. 1–14). This account is contrasted with that of 'certain Pythagoreans', who make matter itself a product of the original Monad, a view rejected as absurd (295. 15–24): thus Numenius is committed to a fundamental duality of principles. Moreover, he criticizes the Stoics for making matter one of the indifferents, rather than assigning to it responsibility for evil, which they are forced to ascribe to 'a sort of perversion', without being able to say what causes it (296–7). Evil must itself be allowed the role of a fundamental principle of things, or its origins will become inexplicable. Numenius' argument implicitly relies upon the principle that effects resemble their causes (PCS: I.3*d*; cf. the Platonic principles P1–4: I.3*d*, III.1*b*). Plotinus will decisively reject such Gnostic or Manichaean dualism (*Enneads* 1. 8, esp. 1–5, 9; 2. 9: §2*a* below).

One other such ontological hierarchy, derived from the first century ad Pythagorean Moderatus of Gades, [9] bears brief mention. It comes from Porphyry's treatise *On Matter*, as preserved by Simplicius:

438 it appears that this view concerning matter was first held among the Greeks by the Pythagoreans, and after them by Plato, as Moderatus recounts: for he says . . . that the First One is beyond being and all essence, while he says that the Second One, which is 'really real' [*ontōs on*: cf. X.3*b*) and intelligible, is the Forms, and the Third, which is Soul, participates in the One and the Forms, while the final nature which comes after it, namely that of perceptible things, does not even participate, but receives order by reflection from the others, perceptible matter being a shadow cast by non-being, manifested primarily as quantity, but of a lesser degree than it. (Simplicius, *On the 'Physics'* 1. 7. 230. 34–231. 5)

This arrangement, in which all of reality is ultimately dependent upon the highest principle, and in which matter simply emerges out of non-being and has no independent ontological role to play, at least prefigures that of Plotinus.[10]
end p.408

2. Plotinus' System

(a) Matter, Form, and Soul

Plotinus' philosophical goal was to create a complete and completely lucid account of reality in which the relations of metaphysical dependence obtaining between its various levels could be exhibited as clearly as the limitations of human language and conceptualization allow. At the head of the entire structure is the One, the single archetypical reality from which, in a manner which Plotinus struggles to clarify, everything else in the universe flows. In so far as his philosophical enterprise is driven by the desire to render the world and its contents fully intelligible, Plotinus' project recalls both Plato and Aristotle; and he makes judicious use of the doctrines of each of them in the construction of his system (cf. §1*a*, **433** above).

He takes over much of the Peripatetic picture of the sublunary world, talking of matter and form in terms which echo (although without exactly reproducing) those of Aristotle (see *Enneads* 1. 8. 10–11; 2. 4; 6. 1. 25). Matter is required to receive form and quality. It is more than merely space (as some, including Aristotle, although not Plotinus himself, understood Plato's receptacle: III.3*d*). Plotinus calls it 'mass', *onkos* (2. 4. 8–12), something intrinsically incorporeal (*contra* the Stoics: 6. 1. 25–8), prior to bodies, and apprehended by the intellect:

439 matter is necessary for both quality and magnitude, and consequently for bodies; and it is not an empty name, but exists as a substrate of some kind, even if it is invisible and without magnitude. Otherwise by the same reasoning we will deny the existence of qualities and magnitude, since each of this sort of thing would be said to be nothing if taken alone and in itself. On the other hand if these things are existent albeit dimly so, all the more so will be matter, even if it is not apparent since it is not apprehended by the senses. . . . but by reasoning. . . . But there is no corporeality about it, since if corporeality is structured, it is other than matter, and matter is something else. But if it has already been productive and, as it were, mixed, it would clearly be body, and not matter alone. (2. 4. 12. 21–38)

Matter is simply privation (2. 4. 13–14); it is unlimited (2. 4. 15), and is intrinsically evil not because it functions ontologically as a separate principle (*contra* Numenius: §1*b* above), but because it lacks the fullness of being associated with the good (2. 4. 16).

Plotinus parts company with Aristotle by emphasizing the importance and distinctness of the intelligible world of forms. However, that world must also involve matter:

440 among the intelligibles, the composite is different and not as it is with bodies. And since the principles (*logoi*) are composite, and by their activity (*energeia*)
end p.409

make composite the nature which acts in the direction of form. . . . The matter of generated things too is always taking on different forms, but that of eternal things is always the same and has the same form. Matter here is pretty much the opposite, since here it is everything in turn, and one at each moment, and so nothing persists, as one thing drives out another. But there it is everything at the same time, and so there is nothing for it to change into, since it already has everything. (2. 4. 3. 6–14)

The reason why there must be such intelligible matter (the concept is ultimately Aristotelian: cf. *Metaphysics* 8. 6. 1045a34–6) is that there is a multiplicity of Forms:

441 if there are many Forms, there must be something common to all of them, as well as some property with which one differs from another. This property and separating difference is each one's proper shape. But if there is shape, there is something shaped, regarding which there is the difference. There is, then, matter receptive of shape and the invariable substrate. Moreover, if there is an intelligible world of which this is an imitation, and this is composite and from matter, there must be matter there as well. (*Enneads* 2. 4. 4. 2–9)

These are Platonic Forms, independently existing, hypostasized, perfect qualities, yet they involve matter (albeit of an intelligible kind), a view which threatens a familiar regress (if the Forms are themselves composite, then there must be some further Form of the Forms in virtue of which they are all of the same basic type; and so on). However, Plotinus is not, I think, motivated by the desire to produce an abstract theory of universals as such; what unites the Forms simply will be the fact that they are Forms (i.e. their causal role in the overall structure of things).

Plotinus argues for the Forms in an early treatise (5. 9), first dismissing Epicurean and Stoic materialisms, and then presenting his own Platonist dualism. Taking his inspiration from the *Phaedrus* and the *Symposium*, Plotinus outlines an intellectual ascent away from the material world and into the realm of pure Mind (*nous*: 5. 9. 2) which involves first recognizing the distinction between soul and body (not only in human beings but in the world as a whole), and then recognizing that soul is itself a composite of matter and form (the form being Mind):

442 then again you will investigate whether the soul is one of the simples, or whether there is in it something like matter and something like form, the mind in it, one part of which is as the shape is to bronze, the other as the producer of shape in the bronze. And transferring these things to the Whole, one will ascend there too to Mind supposing it to be the true producer and Demiurge, and will say that the substrate, in receiving the forms, becomes fire, water, air, and earth, and that these forms come from something else, which is Soul. (5. 9. 3. 21–30)

Thus matter in the sensible world is informed by form, which proceeds from Soul, the lowest of Plotinus' three hypostases, corresponding to

Plato's Demiurge (and to Numenius' Third God: §1*b***434**, **437**, above; cf. Moderatus, **438**).

(b) The Priority of Actuality

Why not, Plotinus then asks (*Enneads* 5. 9. 4), simply stop with Soul, and assume that it will generate Mind on its own? The answer is that Mind, or pure abstract thought (as contrasted with practical thinking, or *dianoia*: cf. Numenius, **437**), is better (in the sense of being more perfect) than Soul, and what is better must be prior in actuality:

443 for certainly Soul, when it is completed, does not produce Mind, as they[11] hold; for how will what is potential become actual if there is no cause to lead it into actuality? For if it is a matter of chance, it is possible for it not be actualized. Thus we must posit the primary things as being actual. (4. 4–8)

Thus Plotinus relies on a Principle of Prior Actuality (PPA) (compare Aristotle, IV.2*a*, V.2*g*; *Physics* 2. 3. 195^{b}4–6; *Generation of Animals* 1. 19. 727^{a}25–30, 2. 4. 739^{b}19–30; *Metaphysics* 7. 8. 1033^{b}32, 9. 8) to evade the unpalatable conclusion that formal structure could simply emerge from chaos. He goes on to argue that, since the soul is subject to affection, then if everything is not to be utterly annihilated in the course of time, there must be something prior to it and unaffectable to guarantee its continued existence (cf. III.2*a*, **131**). Similar arguments are rehearsed in another early treatise, on the immortality of the soul (4. 7. 8^{2}), in which Platonist arguments for the soul's being an incorporeal substance are urged against both Stoic materialism and Aristotle's functional account. Mind, then, is prior to and independent of Soul, and contains (indeed is identical with) the Forms in their truest state. The latter are contents of Mind, being neither prior (as Porphyry at one time held: *Life* 18) nor subsequent to it: ‘mind thinks what is really real, not as if they were somewhere else; for it is neither prior nor posterior to it’ (5. 9. 5). Mind's relation with its contents is not one of subject and object, nor does the existence of the Forms compromise Mind's essential unity, since they are not separate in the sense of being separately located in Mind. Rather ‘Mind is all things together, but also not together, because each of them is a particular capacity’ (5. 9. 6). Each Form is simply a distinct aspect (or potentiality) of Mind. Thus Mind's intellectual activity cannot divide itself into thinking subject and thought object (Plotinus, echoing Aristotle, holds that it thinks of itself). Plotinus struggles to characterize this ‘relationship’:

end p.411

444 but if Thought and what is thought are identical—for Thought is a type of actuality, since it is clearly not a potentiality . . . —then what is thought will be the primary substance. So if it is actuality—and, moreover, the first and fairest actuality—it will be Thought, indeed substantial Thought, since it is the truest. But this sort of Thought, being primary and primarily so, will be the primary Mind, since this Mind is neither potential nor is it one thing and its thought another, since in this way again its substantiality would be potential. So if it is actuality and its substance is actuality, it will be one and identical with its actuality; but what is and what is thought are also one with its actuality; thus it will be one and all of them at the same time, Mind, Thought, and what is thought. (5. 3. 5. 33–44)

Fortunately we need not straighten out the sinuosities of that typically serpentine passage.[12] The driving force behind his desire for such unification is the attempt to apply throughout his system what has been called the Principle of Prior Simplicity[13] (PPS). It is on the basis of PPS that Plotinus both argues for the essential unity of Mind, and yet is forced, in view of the extent to which his Mind is still (potentially at least) multiple, to posit a yet higher and more fundamental, absolutely unified principle, the ineffable One. PPS is expressed in several places:

445 if there is something after the First it must come from it, and have its ascent to it either immediately, or through intermediaries, and there must be an ordering of seconds and thirds, the second ascending to the first and the third to the second. For there must be something simple prior to everything, which must be different from everything subsequent to it, existing on its own, unmixed with its derivatives, yet capable of being present to the others in another way, being really One, and not being many in one way and one in another; and it is false to say concerning it even that it is one, of which 'there is neither account nor understanding', [14] and which indeed is said to be 'beyond being'. For if it is not to be simple, and beyond all coincidence and composition and really One, it would not be a principle (*archē*); and it is most self-sufficient in its being simple and First of everything. For what is not first requires what is prior to it, and what is not simple requires the simples it contains in order to come to be from them. (5. 4. 1. 1–15)

Plato placed the Form of the Good above all the others, describing it as being 'beyond being' (*Republic* 6. 509b; cf. §1*b*, **438** above). Plotinus adopts that phrase and applies it to his One, the single, absolute, transcendent entity that articulates and gives structure to the whole of subsequent reality, both intelligible and sensible. Here Plotinus glances both at

end p.412

Speusippus' neo-Pythagorean Platonism (XI.1*a*), and at Numenius (§1*b* above); but the development of the system is original.

The One must, in Plotinus' view, be utterly simple in order to function as a principle. If unity is simply an attribute of other things (as Aristotle thought: *Metaphysics* 10. 1–2), then it will itself cease properly to be unified, which is an absurdity in Plotinus' view (*Enneads* 5. 6. 3. 10–15). Plotinus is clearly influenced here by the antinomies of Plato's *Parmenides* (137b–143a, from which he quotes in **445**), which tease out the consequences of the proposition that the One exists, or has being. Any predication of anything of the One will compromise its essential unity (which is why it is 'beyond being'; cf. *Enneads* 1. 8. 2; 5. 3. 13, 17; 5. 5. 6; 6. 7. 33, 37). Thus a thesis about the logical status of such concepts as 'one' generates profound metaphysical consequences. Equally, the Plotinian conditions on the One may be seen as flowing from the requirements of maximal intelligibility and the explicability of order which characterize (in various forms) Greek natural teleology, the exposition and analysis of which has occupied much of this book. Plotinus argues that if there is more than one basic principle we must either suppose, *contra hypothesem*, some further unifying principle over and above the multiplicity posited in order to account for their particular combinations which will otherwise be fundamentally inexplicable: but such randomness is intellectually intolerable. Plotinus deploys this argument against Aristotle's cosmology: if all of the heavenly bodies are independent, primary principles, they must merely be a chance collection (5. 1. 9; however, cf. V.2*g*).

Another passage is worth quoting in this context. After arguing for the necessary multiplicity of Mind (5. 3. 9–11; cf. 5. 4, 5. 6), Plotinus writes:

446 so if Mind is because it is multiple, and thought itself, even if it comes from it, is a type of event which makes it multiple, the absolutely simple and First of all must be beyond Mind; for if it is to think, [15] it will not be beyond Mind, but if it is Mind, it will be multiple. But why should it not be multiple in this way, even while it is a single substance? For the multiplicity [sc. of Mind] is one of activities, not composition. But if its activities are not substances but proceed from potentiality to actuality, while it will not be multiple, it will be imperfect prior to actualizing its substance. But if the activity is the substance, and its activity is multiple, its substance will be just as multiple as well. . . . And there must be One before the many, from which the many derive, since in every number-sequence one is first. But in the case of number they[16] say that this is so because the sequence is compositional; but in the case of existing things, what necessity is there that there already be a One in this case too from which

end p.413

the many derive? But then the many will be dispersed apart from one another, each coming haphazardly into composition from one place and another. (5. 3. 11. 26–12. 14)

PPA is invoked to rebut the notion that the One, like Mind, may be perspectivally multiple: if it were it would be incomplete, in the sense that it would involve movement from potentiality to actuality, and, by PPA, such a movement presupposes a higher, prior actuality. Equally, composite things cannot simply be arbitrary agglomerations (this is developed in another treatise which attempts to clarify the nature of the One: 6. 9. 1–2). The insistence on total intelligibility, which in turn entails prior simplicity and actuality, pushes Plotinus inexorably towards his system of three hypostases, with the One standing in magnificent isolation at the top.

But the very uniqueness and simplicity of the One poses serious problems for Plotinus, problems again deriving from the antinomies of *Parmenides*. For if the One really is without attributes or predicates, how can anything be said intelligibly about it? Plotinus frequently wrestles with the ineffability of the One: we can say nothing of it and it cannot be known; we can only ‘try so far as possible to make signs to ourselves about it’ (5. 2. 13. 5–7; cf. 6. 8, 6. 9 and 5. 3. 12–17; and compare Alcinous, *Primer* 10: X.3*d*).

The situation is not, however, hopeless; and Plotinus' attempt to resolve the difficulty provides a convenient bridge into discussion of the causal relations his system posits between the various levels of reality:

447 to call it [i.e. the One] a cause is to predicate some accident not of it but rather of us, because we have something of it while it is in itself. But strictly speaking one must not say ‘it’ or ‘is’, but it is as though we are circling around it on the outside in our desire to explain our own experiences, now coming nearer, now falling away, in the difficulties which surround it. (*Enneads* 6. 9. 3. 49–55)

When we talk of the One as cause we are speaking of its effects in us, rather than of any attribute it possesses (cf. Proclus, §3*b* below).

(c) The Causal Structure of Reality

Plotinus' commitment to fundamental intelligibility and explicability led him to postulate an intellectual Mind beyond the creative Soul, and a One beyond both which is ultimately responsible for everything that there is. But how? The second great challenge for the

Plotinian system is to explain how such a One could have effects of any kind at all. Plotinus sees the difficulties involved:

448 how then does it [i.e. Mind] see, and what, and how does it exist and how has it come to be at all from it [i.e. the One] in order to see? For the soul now knows that these things must be, but desires an answer to the vexed question

of the ancient sages, namely how, from One being such as we say it is, can anything, be it multiplicity, Dyad, or number, come into existence, and why did it not remain in itself, but rather poured forth that great multiplicity which is seen to exist in things but which we yet think should be referred to it? (*Enneads* 5. 1. 6. 1–8)

The One's uniqueness cannot be undermined in any way by the derivation of the rest of reality from it, nor can it undergo any change (thus any implicit principle involving reciprocal action of the effect on the cause must be abandoned). Plotinus sketches his solution in a number of places, making use of a series of images and analogies:

449 so it is necessary that, if there is a Second after the First, since the latter is immobile, it must have come to be without any inclination, wish, or change of any kind on the latter's part. How then, and how should it be conceived in relation to the stationary One? As an illumination from it but one which leaves it unchanged, like the brilliance of the sun which, as it were, circles around it and springs from it continually but leaves it unchanged. And all existing things, as long as they persist, necessarily produce around themselves existence which depends upon themselves, from their own substance towards the outside, from their own present power, as a sort of image of the archetypes which produced it. Fire produces heat from itself, while snow does not preserve its cold only within itself. This is particularly clear in the case of fragrances: for as long as they exist, they emanate something from themselves to their surroundings, whose existence is enjoyed by those around. And everything which is come to perfection generates; and what is always perfect always and eternally generates. (5. 1. 6. 25–39)

This talk of the outpouring or outflowing of being (**448–9**), which owes something to Plato's image of the sun (*Republic* 6. 508b), recurs in the discussion of how Mind, the second hypostasis, brings Soul into existence:

450 Now this [i.e. intellect] being like that [i.e. the One] produces in the same way, pouring forth its great power, which is a form of it, just as it was poured forth by what came before. And this activity out of the substance of Mind is Soul, which comes to be while the former remains unchanged. For Mind comes to be while what is prior to it remains unchanged. But the latter [i.e. Soul] does not remain unchanged when it produces, but by being changed generates an image. (*Enneads* 5. 2. 1. 14–19)

The image generated by Soul is the life-force found in all individual living things, including plants. The idea that the stream of production from the One leaves the One precisely as it is expressed elsewhere in a number of different, although equivalent, ways (cf. 4. 8. 6; 5. 1. 3; cf. Numenius, §1*b*, **435** above; Proclus, §3*b* below).[17] Writing elsewhere about the One, Plotinus asks

451 what then is it? The power (*dunamis*) of everything. If it did not exist, neither would everything else, and nor would Mind, which is the first and whole life. But what is above life is the cause of life, since the activity of life which is everything is not primary, but flows out as it were from a spring; for consider a spring which has no other origin, but which gives itself to all the rivers but is not exhausted by them, but remaining itself unchanged. (3. 8. 10. 1–8)

The One cannot be potential in the way matter is, since that would imply incompleteness; rather it simply generates everything ultimately from itself, without suffering diminution (cf. 6. 9. 5).

That this process of generation is not in any ordinary sense voluntary is implied by **449**, and made explicit elsewhere. *Ennead* 4. 8, which deals with freedom and necessity as they relate both to human beings and to the One and Mind, emphatically denies that the One's activity is constrained in any way, even though what it does is in a sense necessary (6. 8. 7).[18] The One is what it is, but is not forced against its will to be so, since such constraints only operate in things which are essentially divided, multiple, and subject to conflict: there is nothing to which the One can be enslaved, as we may be to the passions (6. 8. 12). We may speak (inaccurately, since such talk itself compromises the One's unity) of an identity between the One's substance and its activities, which we may also (again loosely) ascribe to its will (6. 8. 13): but such talk is no more than a handy heuristic device.

The One, then, is the producer of all activity in everything, but not by a direct process of intervention, in the way in which Soul (with the backing of Mind) produces the visible world:

452 if there is to be something after it [i.e. the One], it can no longer be simple; therefore it will be One–Many. Where does it come from? From the First, for surely not by accident, or else it would not be the principle of everything. So how, then, does it come from the First? If the First is perfect, indeed the most perfect of all, and the primary power, it must be the most powerful of all things, and the other powers imitate it as far as they can. Now, when anything achieves perfection, we see that it generates, and does not allow itself to remain alone unto itself, but makes another; and this is not only the case for things which have choice, but also for whatever grows without choice, and even inanimate things give of themselves in so far as they can. For example, fire warms, snow cools, drugs perform something else; all of them imitate the first principle as far as possible in respect of eternity and goodness. So how could the most perfect good remain within itself, as if it begrudged itself or was incapable, it being the power of all things? How could it still be the principle? Something must come from it if anything is to be from the others whose existence comes from it; that they do come from it is necessary. (5. 4. 1. 21–42)

end p.416

The productive hierarchy is necessary, but not physically so. Rather for conceptual reasons it could not be otherwise than that the One produces in the way it does: that is simply what it is to be the One.[19]

The One generates the rest of reality by a spontaneous emanation of power or ‘procession’ (*proodos*); but equally everything that is produced seeks to imitate the perfection of the One, in the opposite process of reversion (*epistrophē*) towards the One as final cause, whence the other orders of reality derive their structure (compare Proclus, §3*b* below). This imitation involves self-perpetuation to the best of one's ability (cf. Aristotle, V.2*g*), since if the drive to self-perpetuation is universal it must inevitably be ascribable to the One. The following chapter sheds more light on the issue:

453 for each thing, there is one activity of its substance and another from its substance; the activity of the substance is the same as the thing, while the other is from that one, being in each thing a necessary consequence of and distinct from it. For example, in the fire there is the heat which is constitutive of its substance, and the other which is generated by it when it acts as is natural to its substance as enduring fire. Things are the same in the higher realm; and all the more so there, where it 'remains in its proper life', while the activity generated by its own perfection and its concomitant activity takes on substantial existence, since it derives from a great power, indeed the greatest of all, and attains being and substance. (5. 4. 2. 27–41)

Thus there are two activities involved, one proper to the internal nature of the thing, and which indeed constitutes it, the other which it induces in the external world. The necessity of this external induction is ascribed directly to the Good (2. 9. 3): it is of the nature of the Good to give of itself. Thus the whole of reality consists in a descending ladder whereby higher entities, in virtue of their external activities, create the characteristic internal activities of beings on the rung below; indeed, the externally directed activity of the higher entity actually becomes the lower (5. 4. 2). These activities are, even at the lowest level, properly speaking immaterial:

454 but if, when they see the activities of bodies heating and cooling and pushing and pressing down, they rank the soul with them, setting it up in a sort of active field, they are first of all ignorant of the fact that even these bodies form these things by virtue of the incorporeal powers within them. (4. 7. 8^{1}. 1–6)

The argument for the existence of incorporeal powers (which would not have swayed any Stoic: VII.1*a–b*) is that quality is distinct from body, in a way that quantity is not; and hence, since quality is not part of the essence

end p.417

of body as such, qualities must be incorporeal. We shall return to this distinctive strand of Neoplatonic immaterialism below (§3*b*).

We may conclude this section with one more extended passage that bears directly on Plotinus' concepts of cause and explanation, and their relation to substance. Plotinus begins his treatise on the forms and the Good (6. 7) by stressing that the generation of the sensible world, which is atemporal (§ *d* below), involves no deliberation or planning, since they are temporal concepts. The question then arises of whether we can assign any reason or cause to the higher hypostases, since they are not the products of free creation:

455 we cannot thus see how great a thing Mind is: for we allow it to have a 'that' but not a 'why', or if we do, as something distinct from it. And we see a man or an eye, as it might be, as an image, or something of an image. But there [i.e. in the intelligible realm] there exists man and the reason why there is a man, if at least the man there must be something intelligible, and the same for the eye and the reason why: they would not be there at all if there were no reason why. But here, just as each of the parts is separate, so too are their reasons. But there everything is in one, so that the thing and its reason are identical. Even here things are frequently identical with their reasons, as in the case of an eclipse.[21] So what prevents everything being its own reason in other cases as well, this being the substance [or essence: *ousia*] of each thing? In fact this is necessary, and when we attempt to grasp what it is to be something, things turn out right, since what each thing is is its reason. I do not mean that each thing's form is its cause—although it is—but that if you unpack each thing's form in relation to itself you will find the reason within it. For while something inert and lifeless has no form at all, where could the form which belongs to Mind get its reason from? If someone says 'from Mind', then it is not separate from it, since it is Mind; so if it must possess these things in no way lacking, it will not lack the reason. But Mind possesses in this way every reason for the things in it, but it is also all of the things in it, so that none of them requires in addition any reason for their generation, but it has come to be everywhere and has in itself the explanation of its existence. . . . So it also gives to things which share in it in such a way that they possess a reason. And just as in the totality here which is composed of many interrelated things, whose individual reasons are contained in their being all, and just as in each individual the part is seen in relation to the whole, it is not that this comes to be and then that after it, but they each set up cause and effect jointly in relation to one another, so much more must it be the case there that each thing is related to the whole and each to itself. (6. 7. 2. 2–37)

Characteristic obscurities aside, the gist of this is clear enough. The distinction between reason and object, cause and effect, can only be made in the case of parts which make up a whole (and which exist in temporal sequence):

end p.418

for the whole itself there is no such Aristotelian separation of reason and fact (cf. Aristotle, *Posterior Analytics* 1. 13; V.1*b*, **179**).

This is not without its problems, however. Plotinus' subject here is Mind, and it is unclear how far, if at all, he would have wished to extend the analysis appropriate to mind to the other levels of reality; yet the talk of parts and wholes seems perfectly general. Moreover, it is hard to square what he says in **455** about the self-explanatory nature of the whole Mind with his explicit remarks elsewhere to the effect that the One is the cause of Mind (6. 7. 16), while an insistence on the self-explanatory nature of even one of the lower hypostases apparently undercuts the argument to the existence of the One based on the Principle of Prior Simplicity (§ *b* above). The key, I think, is the Neoplatonist insistence of the immanence of causes in their effects (indeed, if anything, the relation is the other way round: effects subsist in their causes; cf. Proclus, §3*b* below). Taken as a whole, Mind is indeed a self-explanatory unity: but only because of its participation in the One. Of course it is a further question just how intelligible those claims can be made.

(d) Cosmology and Freedom

A similar insistence on the inseparability of reality from its explanation informs Plotinus' discussion of the creative organization of the world. His treatise *On the Heaven* begins as follows:

456 if, when we say that the world has always existed and will always exist even though it is bodily, we adduce as the explanation the will of God, we may very well be speaking the truth, but we will be offering no clarity of explanation. (*Enneads* 2. 1. 1. 1–4)

We may suppose, on the grounds that things in the material world are observed to be in continual reciprocal flux, that God operates in the same way on everything, and thus assume even the heavenly bodies to be mutable. But if we wish to insist on the immutability of the heavenly bodies (rather than simply upon some celestial conservation principle) we need to show in addition that the divine will extends to the preservation of the celestial individuals as the individuals they are in virtue of their souls (since Plotinus rejects Aristotle's ether: 2. 1. 2. 11–14; cf. Philoponus, §3*c* below). Plotinus' solution looks back to the *Timaeus* (37c–d); given that it has been good for there to be a world organized the way it is, it is unthinkable that any such good should be subject to flux:

457 how can anything which has once been placed in Soul, which is moved by a remarkable power and is set next to the best things, ever escape from it into non-existence? Not to think that Soul, which arises from God, is greater than any bond is the mark of a man unversed in the reason why everything holds

together. For it would be absurd for Soul, if it can hold it together for any length of time, not to do it for ever, as though this holding together were forcible and its natural condition were other than this. (*Enneads* 2. 1. 4. 14–21)

The mistake identified by Plotinus (in implicit rejection of Aristotle's cosmological arguments) is to suppose that something that happens by a certain necessity of soul is necessarily thereby unnatural. This rejection of the Aristotelian notion that the heavens move as a result of their consisting of a peculiar essence leads Plotinus to a novel account of the celestial circular motion. It is neither wholly natural (i.e. due to the material), nor wholly due to the influence of Soul, but is rather the sum of the two tendencies. Matter tends to move in a straight line, while Soul tends to pull inwards, and the result is perfect circular motion (2. 2. 1). This remarkable doctrine of the heavenly motions resulting from a certain dynamic tension obviously (if crudely) prefigures the physics of Descartes and Newton (compare Plutarch, X.3*c*, **371**).

The relations between souls and matter on the one hand and the higher hypostases on the other are a matter of obvious concern to Plotinus. On the one hand, he seeks to show that, although embodied souls are naturally constrained by the material nature of their embodiment, this need not fatally compromise their freedom. But equally, if he is to prise humans free from the grip of cosmic determinism, he must show how it is possible for individual souls to be free in a universe which is (for him) both providential and ultimately causally dependent upon the overflowing power of the One as it ramifies through the lower hypostases. And he must further clarify both how matter can be intrinsically evil without compromising the unsullied goodness of the One, and how the One, being whole and indivisible, can in fact be present to all the rest of creation.

Plotinus begins his early treatise *On Fate* (*Enneads* 3. 1) by making a disjunctive classification of the formal possibilities available as regards the extent of causal explanation:

458 everything that is generated and everything that exists is either generated (if it is generated), or exists (if exists), in accordance with a cause; or both are without a cause; or some are caused and others causeless in each category; or everything generated is caused, while of the things that exist some are caused and others are not; or conversely everything which exists is caused, while of the things generated some are caused and others are not; or none of them is caused. (3. 1. 1. 1–8)

Plotinus then deals with each of the alternatives thus enumerated. The primary eternal things cannot have a cause, or else they would not be first (in explaining the structure of their activity it is enough to refer to their essences); conversely, all things dependent upon them, just in so far as they

are so dependent, are caused by them (Plotinus here implicitly allows that there may be more than one genus of causation). On the other hand, in the case of all generated things, we must posit individual and exhaustive causes, since randomness of the Epicurean type is intolerable (VI.3*a*; cf. VII.1*a*; Plotinus thus accepts CP1*a*, that nothing occurs causelessly: I. 2*d*; VI.2*a*); and Plotinus is equally clear that freedom cannot rest upon 'senseless impulse' (3. 1. 1. 8–24).

Causal ascription is relatively easy, he holds, where the causes in question are not remote (he offers a number of examples drawn largely from Aristotle: 3. 1. 1. 24–35). None the less, the genuine philosopher must not stop at such immediate causes, but must attempt to understand the overall causal structure of things. Plotinus sketches the views of Epicureans, Stoics, and astrological determinists (3. 1. 2), criticizing each in turn (3. 1. 3–7) before outlining the proper Platonist account of the matter (3. 1. 8–10). The Epicurean view involves the absurdity of order emerging from chaos (VI.1*d*, 2*a*, *c*; cf. III.4*d*). On the other hand strict determinists with their adamantine chain of causation abolish all secondary causes (and hence individual responsibility), since we might as well say that it is the immanent ruling principle itself that actually does everything (3. 1. 4). Plotinus seeks to show that, on determinist suppositions, human 'agents' are really merely instruments of the overall causal structure of things, and hence not really agents at all (he compares human 'agency' in this case to that of a leg: 3. 1. 4. 12–24; cf. XI.4*b*). Plotinus detects no significant difference between this account of universal direct determination by the immanent *logos* and Chrysippus' ostensibly more sophisticated story in which individuals are assigned causal status, but one which is itself determined by the sum total of previous events (3. 1. 7: cf. VII.2*a*–*c*, e; X.4*b*–*c*).

Equally, Plotinus rejects astrological determinism of the sort which suggests that all human events and actions are the exclusive result of the positions and influences of the heavenly bodies (3. 1. 5–6). The position rejected here is that attacked by Sextus (VIII.5*c*); but, while Plotinus does not deny that some influences percolate down to earth from the heavens (cf. *On Whether the Stars are Causes*: 2. 3. 11–12), he repudiates any version of 'hard' astrology, even of Ptolemy's sophisticated type (XI.1*b*).

The reasons for this are rehearsed at length in *On Whether the Stars are Causes*. The stars, being divine, cannot be directly responsible for evil, as the astrologers hold, nor can their mere relative positions cause their particular dispositions, since they are ensouled and hence rational. Moreover, the astrologers allow too much individuality to the planets, failing to see them as simply bit-players in the great cosmic drama directed by the First Principle (2. 3. 1–6; Plotinus uses the dramatic metaphor elsewhere: 3. 2. 15–18). On the other hand, it is reasonable to assume, on both empirical and general conceptual grounds, that they may function as

end p.421

general dispositional signs (2. 3. 1, 3. 1. 5), and this can be explained in terms of the general harmony of the whole (2. 3. 7–8). The stars, then, signify our general dispositions; but

459 we are what we really are, we whom nature has endowed with the ability to overcome our passions. (2. 3. 9. 15–16)

The stars signify without genuinely acting (they do so only in so far as they are 'affections of the totality': 2. 3. 10. 1–3; cf. 4. 4. 30–1), which is why diviners can tell us what will happen, but not why: they simply read what is written into nature (3. 3. 6). And even if the stars do exert a direct causal influence upon our characters, they cannot be held responsible for our defects, which are rather due to a failure to emulate divine perfection (2. 3. 11).

Even so, Plotinus is aware that the problem of evil must be confronted. He is implacably opposed to the Gnostic view that evil must be posited as a separate principle identified with matter (2. 9); rather matter itself is a falling short of divine perfection, and the evil it embodies purely and simply privation (2. 4. 16; §2*a* above). But if matter itself is part (albeit the lowest one) of the successive outpourings of being from the One, how is the One not responsible for evil? Equally, providence (to which he devotes a long treatise: 3. 2–3) cannot be the direct cause of the evil that men do. Yet the Principle of Prior Actuality (PPA: § *b* above) apparently implies that evil must have a prior, higher cause. Plotinus struggles to reconcile these apparently conflicting doctrines:

460 now let us say how it is that we say that 'the totality is under the rational direction of Soul'.[22] Does it make each thing as it were in a linear fashion: men, then horses and other animals, then wild beasts, then fire and earth primarily, and then seeing these come together to destroy and support each other, seeing only the interweaving from them and the eternal subsequent generation of their consequences, making no further contribution to what follows, but only producing the initial generations of animals and consigning them to their mutual interactive affections? Or do we rather say that Soul is the cause of things generated in this way too, since they produce their subsequent effects having been generated by it? Or again, does the rational principle (*logos*) include this particular thing doing or suffering this, so that not even this sort of event occurs at random or haphazardly, but rather thus necessarily? Is it then the rational principles that produce these things? Rather, although the rational principles exist, they do so as knowing, not as producing—or better the soul which encompasses the generative principles knows everything that will come to be as a result of its actions. For when the same circumstances coincide, it is at all events fitting that the same things should result. Soul takes on and foresees these things, and accomplishes their consequences and

weaves them together, so that antecedents and consequents [are linked], and are further themselves in these cases antecedents of what follows. (2. 3. 16. 5–26)

Plotinus seeks to carve out a non-trivial sense in which Soul may be the cause of everything, even though it is not responsible for every outcome (compare pseudo-Plutarch, X.3*c*, **374**; and contrast Alexander, X.4*b*–*c*). This is, he concedes, difficult to swallow; but the alternative is, he supposes, even worse: the good rational principles must themselves contain the seeds of evil (2. 3. 16. 36–41).

Plotinus makes some play with an argument beloved of all proponents of theodicy, namely that the existence of particular evils in the world is actually conducive to its overall goodness (2. 3. 16. 42–54), a theme developed at greater length elsewhere (3. 3. 5, 7; cf. 3. 2. 5, 13–14; compare Descartes, *Meditations* 4; Leibniz, *Discourse on Metaphysics* 1–3; *Theodicy*; and cf. VII.2*e*, 3*a*). If there were no evil, he thinks, the totality could not itself be perfect (2. 3. 18), presumably because it is of the nature of the perfect to produce, yet what it produces must necessarily fall short of itself in perfection. The visible universe, which, being multiple, necessarily contains conflicts,

461 came to be not because of a calculation that it ought to be generated, but because it was necessary that there should be a second nature: for that [i.e. the intelligible universe] was not such as to be the last of realities, since it was first, and was possessed of much, indeed all, power, this being that of producing without seeking to produce. (*Enneads* 3. 2. 2. 8–13; cf. 3. 3. 3)

This necessity of production makes it otiose to blame Intellect for having created an inferior world; moreover, even if its parts taken serially may be imperfect, they are still contributory parts of a perfect whole (3. 2. 3; cf. 15; and compare VII.3*a*). Thus Intellect cannot be held responsible for the continual strife, destruction, and ephemerality of the perceptible world. Indeed

462 the injustices men do to one another may have their cause in their striving towards the good: when they fall short through their incapacity to achieve it, they turn on each other. (3. 2. 4. 20–3)

Plotinus again insists that order cannot arise from disorder (3. 2. 4. 25–31). The causal dependency must be the other way around, even though it is wrong to say that the orderly actually produces disorder (3. 2. 4. 26–36):

463 living things which have the power of controlled self-movement will incline now towards the better, now towards the worse; and perhaps it is not right to inquire into this self-generated turn towards the worse, since a slight initial turning may become greatly amplified as it goes along, making the fault continually greater and more all-embracing; and lust necessarily coexists with the body. (3. 2. 4. 36–41)

end p.423

That latter disclaimer is fairly unsatisfactory; and elsewhere Plotinus attempts to offer a richer account of the necessary (yet none the less good) generation of evil. Individual souls, if they are to be capable of self- movement (and hence properly souls at all), must have a certain freedom of choice, 'since it is not right for providence to exist in such a way as to make us nothing' (3. 2. 9. 1–2; cf. 3. 3. 7). And the fact that for men to do wrong is in a sense in accordance with providence implies neither that providence causes their actions, nor that everything is, in an absolute sense, good (3. 2. 10–11). Evil actions

464 are interwoven [sc. with the good], but are not produced by providence. Rather, if anything useful follows from what has come to be either as a result of men or as a result of something else, be it animate or inanimate, it is reabsorbed into providence in such a way that excellence is everywhere dominant, and the errors are changed and receive correction, just as in an individual body to which health is granted by the providence of the animal, when a cut or wound of some sort occurs, the directing rational principle brings it together again and closes it, and heals and restores the injured part. Consequently, evils are consequences, and necessary ones; however, they come from us as their causes, and not compelled by providence, but rather we bring them together with the direct or indirect works of providence. (3. 3. 5. 24–37)

What we do, Plotinus continues, even when it is good, is done by us and not by providence, although it is done in accordance with providence; while the evil that we do is not even done (at least directly) in accordance with providence (3. 3. 5. 46–9).

The upshot is that we are created with freedom of choice, because it is better and indicative of greater power to create individuals who are themselves autonomous agents than simply to produce a sequence of blind causal consequence. But that freedom, in concert with the fact that in our existence in the perceptible world we are necessarily enmattered, entails that we will from time to time choose evil. However, for Plotinus, as for the Stoics, Galen, and the Christian tradition, real freedom involves transcending the evil influences of the body: it is freedom from the passions. In fact it consists in approximating oneself as far as possible to the divine condition of disembodiment, of transcending the limitations of the corporeal. These issues, as well as the senses in which the One may be said to be free even though its operations are necessary, are followed out in detail in the treatise *On the Voluntary and the Will of the One* (6. 8).[23]

Thus Plotinus has a place in the long and tortuous tradition that seeks to justify the ways of God to Man. But the end-result is of doubtful coherence; and Plotinus is the prisoner of his own rigorous insistence on the
end p.424

absolute primacy of order, and the necessity of the maximal intelligibility of the world, an insistence expressed in his commitment to the principles I labelled PPA and PPS (§ *b* above). This drives him to conceive of the generation of the world and its contents, including even matter itself, as the last stage in a spontaneous outpouring of potency from a One which must itself nevertheless remain perfectly unified, its presence in other things not being a matter of its being parcelled out among them, but rather simply a consequence of their looking back towards it, a doctrine worked out in systematic detail by Proclus (§3*b* below). But in the final analysis, that outpouring itself remains mysterious, and in a sense unexplained. Plotinus' fundamental worry is not Heidegger's, of why there is something rather than nothing, but rather that of why there are many things rather than simply One.

3. After Plotinus

(a) Iamblichus

After his death Plotinus' doctrines were kept alive by his disciples Amelius (§1*a*, **432** above) and Porphyry, albeit with some significant alterations and amendments. Amelius divided the Intellect into three parts (Proclus, *On the 'Timaeus'* 1. 306: cf. Numenius, §1*b*, **435–6**), and posited an infinity of Forms (Syrianus, *On the 'Metaphysics'* 147), including one for Evil. Porphyry's own philosophy took the form of an increasingly scholastic systematization of the elements of Plotinian thought, a trend that finds its highest expression in Proclus (§ *b* below). Porphyry played down the metaphysical distinctions between the hypostases, and (if he is the author of the anonymous commentary on the *Parmenides*) made the three of them (Being, Life, and Intelligence) part of a single, self-unifying principle of intelligible reality: there is thus no separate, prior One in this system.[24] He was also hostile to theurgy (the attempt to achieve mystical transcendence by the practice of ritual) since it involved the impious belief that human actions could coerce the gods (*Letter to Anebo* 1. 2c, 2. 3a–b, 8a-c, 13b, etc.);[25] and he rejected divination as well (ibid. 2. 7).
end p.425

His most important pupil, Iamblichus (fl. *c*.300), however, wrote a lengthy treatise *On the Mysteries* in defence of theurgic practices. He insists that only the correct performance of theurgic ritual can effect the true union with the divine: philosophy on its own is powerless to achieve such an end. In his view, Porphyry wrongly ascribed the successes of theurgic practice to cosmic sympathy, which is merely a *sunaition* for the true cause, divine grace (*On the Mysteries* 2. 11). Iamblichus also insisted, against Porphyry, on the supreme transcendence of the One; and if Damascius is right (*On Principles* 1. 86), he posited a principle higher even than the One, named simply (if paradoxically) the Ineffable, since unity is itself a relative concept and can thus have no place in perfect, undivided reality.

In sharply contrasting style, Iamblichus also composed *On the General Mathematical Science*, which outlines a remarkably sophisticated account of the role of mathematics in science.[26] The third chapter of the treatise (which consists of an exposition of 'Pythagorean' doctrine concerning the relations that hold between number and other branches of knowledge) sets out to distinguish the principles of mathematical knowledge in general from those of the other sciences, and to show how they provide a general and common basis for all of them. The basic principles are the limited and the infinite (*General Mathematical Science* 3. 12. 22–5 Festa[27]), but these are different in the case of intelligible and matterless forms (being in these cases intrinsically indivisible) than in that of the actual numbers, where they are principles of division and distinction, and as such cannot be partless (ibid. 3. 12. 25–13. 9 Festa):

465 these principles (*archai*) differ from those which exist in Intellect in that they provide from their own nature the reason for division, multiplicity, size, and composition, while they are separated from those of nature and of the rational principles (*logoi*) of the soul in that they are unchangeable, and because of those incorporeal intermediates which are ranged in between they exist in themselves separated from matter, while the others are in contact with matter. (3. 13. 21–8 Festa)

The next chapter turns to more particular scientific principles (much of this is now ascribed to Speusippus: cf. X.1*a* n. 5, **350**):

466 let there be posited two primary and supreme principles of mathematical number: the One (which we should not even call something which is in any way, because it is simple and because it provides the principle for things which are, while a principle is in no way of such a kind as those things of which it is the principle), and another principle of multiplicity, which is able of itself to

provide division, and which we can show as far as is in our power to be for this reason much like a moist and malleable material. From these, the principles of unity and multiplicity, is fashioned the first genus, that of the numbers which are composed, with a certain persuasive necessity, of both of them. (4. 15. 6–17 Speusippus Fr. 72 IP)

This is fairly standard fare; and Iamblichus goes on to discuss the nature and understanding of mathematics and its relation to the rest of reality. His primary concern is to emphasize, in concert with the Platonico-Pythagorean tradition and against the views of Aristotle and the Peripatetics, the fundamental priority of mathematical objects over those of the sensible world (ibid. 8). However, towards the end he offers some thoughts on the relation of mathematics to physical reality, ideas which, while having their roots in genuine Platonism (cf. III.4*b*, *d*, 5), seem both strikingly original and sophisticated:

467 so for all these reasons the Pythagoreans reasonably used to honour mathematical zeal, and applied it subtly to the study of the universe. . . . They established what was possible and what impossible in the construction of the universe on the basis of what was possible and what impossible in mathematics, and they worked out the celestial revolutions causally in accordance with proportional numbers, determining the celestial measures in accordance with certain mathematical principles, establishing a predictive science of nature on the basis of mathematics, and in general by setting up mathematics as principles as it were for everything else which can be worked out in the universe. (*General Mathematical Science* 23. 73. 17–74. 1 Festa)

Iamblichus accepts Aristotle's idea that the methods of science must vary with their subject-matter, and that the same sort of causal explanation should not be sought in every case (27. 86. 6–87. 16 Festa; cf. Aristotle, *Nicomachean Ethics* 1. 3. 1094^{b}12–14). But his vision of a physical science based upon ontologically prior mathematical realities is fundamentally un-Aristotelian. Having emphasized the different ways in which mathematical methods can apply to different subject-matters (*General Mathematical Science* 31), he remarks:

468 it is sometimes the habit in mathematical science to investigate perceptible things mathematically, for example the four elements by means of geometry, arithmetic, or harmonics, and similarly with the others. For since mathematical science is prior in nature and takes its start from things which are prior in nature, it constructs its demonstrative proofs on the basis of prior causes. (ibid. 32. 93. 11–18 Festa)

Iamblichus then enumerates a variety of such mathematical forms of explanation, [28] including abstraction (the Aristotelian conceptual stripping

end p.427

away of form from matter); unification (the application of mathematical notions to physical objects); representation (treating what is regular in physics in such a way that it can be assimilated to pure mathematics); and reflection (emphasizing the way in which physical objects embody mathematical principles). Finally, there is that

469 by way of explanation on the basis of prior objects, when, having set up in advance mathematical causes, we examine along with them how objects in the perceptible world are generated from them. Thus I believe we may investigate everything in the realm of nature and generation mathematically. And so for this reason much of mathematics is not confined to pure mathematics, but is drawn towards those things which depend upon them, both because of the judgement of those who apply them, but also because of the kinship of perceptible things with them. (32. 94. 7–16 Festa)

Although Iamblichus looks to Pythagoras for inspiration, this is no mere number-mysticism: rather it is an impressive sketch for the methods of applying mathematics to the physical sciences, albeit one whose recommendations were not (and perhaps could not be) followed out for more than a millennium.[29]

Two further aspects of Iamblichus' thoughts about physical explanation deserve brief comment. Simplicius, in his *Corollary on Place* (to be found in his commentary *On the 'Physics'*), reports that 'the divine Iamblichus' (disagreeing with Theophrastus: V.3*c*),

470 in chapter two of the fifth book of his commentaries on the *Timaeus*, writes as follows: 'every body *qua* body exists in place; place therefore co-subsists in relation to bodies and is in no way separated from them. . . . So those who do not make place akin to cause, and relegate it to the status of the limits of surfaces, or empty spaces, or extensions of whatever sort, are at one and the same time both introducing alien concepts and mistaking the whole intention of the *Timaeus*, which was to link nature invariably with creation; thus we should view place as deriving from a cause, just as bodies are primarily related by kinship to cause.' (Simplicius, *On the 'Physics'* (*Corollary on Place*) 639. 24–34)

Place, for Iamblichus, has direct causal power; indeed, it is his equivalent to the original cohesive Stoic containing cause (cf. VII.1*b*, **251–2**; for a different sense in which place might be a cause, see XI.4*c*, **404**; cf. *b*, **402**(iii–iv)).

Finally, Iamblichus rejects the Aristotelian notion (*Physics* 7. 5) that all causing is by contact. His main target is in fact the Stoics, who, while believing in a universal cosmic *sumpatheia*, considered that it was transmitted through the dynamic tension of the material *pneuma* (VII.1*b*). Iamblichus simply denies[30] that any such physical transmission is necessary:

end p.428

471 it is not the case that if some action is of this kind that all must be, nor must it be from the ultimate forms of action (by which I mean pushing and pulling), nor should we concur with the Stoics . . . that the agent produces by way of contiguity and contact. It would be better to say that not everything acts by way of contiguity and contact, but that action occurs in respect of the suitability[31] of the agent in regard to patient, and that many things act without contact, as we all know, and that even in those cases where we see that action requires close proximity, the contact is only accidental, because the things which are involved in the process of action and affection must be located somewhere. . . . In cases where distance between bodies does not impede action and affection and the reception of the activity of the active part, this occurs instantaneously and without hindrance, as for example when the strings of a lyre resonate although they are far from each other, [32] and naphtha ignites although at a distance from the fire. And many things which are in contact do not act at all, for instance plaster or drugs applied to stones. (Iamblichus, in Simplicius, *On the 'Categories'* 9. 302. 28–303. 6; cf. *On the Mysteries* 5. 7).

Contact, then, is neither necessary nor sufficient for the transmission of causal power (compare the view of Galen: XI.5*b*, **426–8**). Iamblichus does not actually rule out some such theory as that of pneumatic transmission: but he regards it as unnecessary, and not entailed by conceptual considerations. There are no reasons a priori why all causation must take place by contact. Given that, the next step is to investigate what sorts of things are prone to affect each other in order to provide empirical flesh for the skeletal notion of mutual causal suitabilities (again compare Galen: XI.2*b*, **397**). Iamblichus' acceptance of the possibility of non-material, instantaneous causal transmission is not, then, merely the product of an increasingly uncritical Platonizing mysticism; rather it is his attempt to free himself from one dogma of causal materialism.

(b) Proclus

But immaterialism is indeed the spirit of the age; and this can be seen most clearly in the work of Proclus (412–85), whose systematic adaptation of Plotinian thought in his *Elements of Theology* owes much to Iamblichus[33] (although he significantly rejected the Iamblichean innovation of the Ineffable principle). Iamblichus' influence may be discerned in the Law of Mean Terms[34] (LMT), which states that between two distinct types of

end p.429

reality there must be an intermediate stage sharing in some of the characteristics of each (*Elements*, Prop. 28; cf. Props. 14, 40, 55, 63, 64, 132, 166, 181).

At the beginning of *Elements*, Proclus employs LMT in distinguishing the One from the many. There can be no such thing as plurality without unity, since there would then be no way of limiting the things which exist (Prop. 1; *Platonic Theology* 2. 1 also argues that such a pure plurality would be simply unintelligible). Thus existence itself presupposes some degree of unity, but the things which acquire their unity in this way by participation are in a sense one, but equally in a sense not one (*Elements*, Prop. 2); but then the One must be prior to all plurality, since were the relation of dependence the other way about, the plurality could not be a plurality at all (Prop. 5). Finally (Prop. 6) Proclus distinguishes two ways in which the constituents of a plurality may be unified, either by participation (i.e. by being members of a unified class), or by being units, the 'primary participators' in the One;[35] of course, consistently with standard Neoplatonic doctrine, this participation relation cannot infect the One itself with plurality.

Elements 7–13 deals with the nature of causes, beginning with the claim that every productive cause is greater than what it produces (compare Descartes, *Meditations* 3; cf. Aristotle, *Metaphysics* 2. 1. 993^{b}24–6; IV.1*b*). Proclus argues for this by elimination. What is produced will itself either be (*a*) productive, or (*b*) sterile; if (*b*), it is obviously inferior in power to whatever non-sterile object has produced it. But if (*a*), then it is either (i) of equal power, or (ii) of lesser power than its cause. But if (i), then it will be just as productive as what produces it, and hence whatever *it* causes will be equal to it (otherwise we would have an inexplicable decline in causal power), and hence everything will be causally equal, which is impossible. Hence the product must be of lesser power than what produces it; thus Proclus offers a proof for Plotinus' Principle of Prior Actuality (§2*b* above), as well as of the Cartesian notion that the cause is greater than the effect (cf. IV.1*b*).

Beginning from the Greek commonplace that everything desires the good (cf. Plato, *Philebus* 20d; Aristotle, *Nicomachean Ethics* 1. 1. 1094^{a}1–3), Proclus next argues that none of the desiring things can themselves be wholly good, since desire is a drive for something else and hence is caused by an incompleteness (*Elements*, Prop. 8); thus the good which such things possess is immanent, and not the transcendental Good. This reflects the Middle Platonist separation between paradigmatic and formal causes

end p.430

(X.1*b*, **354**; 2*d*, **365**; 3*d*, **378**), while ultimately deriving from the distinction between Forms and their instantiations in *Phaedo* (100a–102e: III.1*b*) and *Timaeus* 52a (III.4*c*: **150**; cf. **472**). Properly speaking the Final Cause (the Good) is both unified and utterly transcendent.

Proclus turns in Proposition 11 to efficient causes, proving that everything must derive ultimately from a single, primary, efficient cause, since otherwise everything would be uncaused (at least in the ultimate sense), or causation would be either circular or proceed *ad infinitum* (compare Aristotle, *Posterior Analytics* 1. 2; cf. V.1*a*; VIII.4). Aristotle's argument (which is designed to show that there must be indemonstrable first principles) is, however, compatible with there being any number of them; Proclus, relying in characteristic Neoplatonic fashion on the Principle of Prior Simplicity (§2*b* above), thinks that the possibility of these being a number of disparate causal principles not unified by any higher principle is, literally, unthinkable. Proposition 12 identifies this first productive cause with the Good, on the grounds that there cannot be anything superior to but yet productive of the Good, as (by Prop. 9) any prior causal principle must be. Finally (Prop. 13) the Good and the One are identified.

Propositions 14–24 make use of similar principles in establishing the ordering and relations between the grades of reality. Proclus first argues that there must be something unmoved (as Aristotle does: V.2*g*) and that there must be a self-mover (*à la* Plato: III.2*a*, **131**). If there is no unmoved mover, then everything must owe its motion to something else, which entails either circularity or regress (Prop. 14), both of which are impossible given the existence of an ultimate principle (Props. 1–5) and the Cartesian principle that the cause is greater than the effect (Prop. 7). This in turn entails the existence of a self-mover, since the first thing to start moving cannot (evidently) be unmoved, but nor can it have its motion directly caused by something else, since that something would have to be, *contra hypothesem*, already moving.

These self-movers are souls, intermediate between bodies and the higher orders of reality. That bodies are not of themselves self-movers is evident from the fact that not all bodies move; moreover, Proclus argues (Props. 15–17), bodies are incapable of the type of self-reversion (the act of contemplating oneself) that is characteristic of minds (here again the Neoplatonists make selective use of Aristotle, in particular his dark and controversial doctrine of the active intellect: *On the Soul* 3. 5). But, as Proposition 18 establishes, all things which simply by their existence bestow some property upon others themselves possess that property primarily (i.e. non-derivatively and universally), and in a higher degree (Prop. 7) than the recipients of it. Thus beyond body is Soul (which is responsible for its movement), while beyond Soul is its unmoved principle, Mind, which is itself posterior to the One (Prop. 20).

Proclus then distinguishes between the plurality of things in each order of reality, and the monadic, unifying principle that makes it such an order, and without which it would not be ordered at all: this is the fundamental prior cause of each order as a whole, although the individual elements within the order may themselves manifest subordinate causal relations (Prop. 21). Finally in this section Proclus introduces a further threefold division, one which also manifests LMT, between things in which there is no participation (‘the unparticipated’), [36] those in which other things participate (‘the participated’), and those other things which participate in them (‘the participants’). The unparticipated (the genuine monadic Form) exists at the top of the hierarchy; but rather than remaining sterile, it gives something of itself to the participated things (the individual forms in the objects), which are themselves prior to the actual things of which they are the forms (Props. 23–4). The Form is more closely akin to the cause, the forms less so:

> **472** so the unparticipated is prior to the participated, and these to the participants. For, in brief, the first is one prior to many, the participated is among the many, and is one yet not one, while every participant is not one yet one. (Proclus, *Elements of Theology* 24. 28. 17–20 Dodds[37])

Thus Proclus seeks to resolve the fundamental problem of the simultaneous transcendence and immanence of the Forms; and brings to completion the Middle Platonist adoption of the immanent Aristotelian form to complement its transcendent Platonic homonym (cf. X.2*d*, **365**; the formulae of **472** may be compared with *Platonic Theology* 1. 11. 25, which equates the First with One, Mind with One–Many, Soul with One-and- Many, and Body with Many-and-One).

Propositions 25–39 deal with procession and reversion (§2*b* above), the relations that hold the various levels of reality together. Procession is simply the spontaneous outpouring of being which each echelon bestows on that immediately below it without thereby suffering any decrease in its own power (Props. 25–7: cf. §2*c*, **449–51** above). Things become productive causes out of their own completeness and abundance of power (Prop. 27), producing first in sequence things like themselves, and then things unlike (Prop. 28), since the productive cause cannot produce things exactly like itself (by Prop. 7), nor yet things completely unlike itself, since then there would be no community between cause and effect, and the Principle of Causal Synonymy (PCS) would be violated. Thus once more LMT comes into operation; and there will be a continuity of procession from the ultimate cause down to the lowest material grade of

reality, each stage in the sequence preserving something of the previous one (Props. 29–30).

Yet everything is desirous of the Good, and seeks to revert towards it, by way of whatever intermediary causes intervene between it and the ultimate reality (Props. 31–4); thus

> **473** every effect both rests in its cause and proceeds from it and reverts towards it. For if it should only rest, it will not differ from its cause and be indistinguishable from it, since distinction involves procession. But if it only proceeds, it will have no contact or sympathy with its cause, having no sort of connection with it. But if it only reverts, how could that which has no substance from it make a reversion in regard to substance to something foreign? (35. 38. 9–15 Dodds)

And Proclus proceeds to show that the other possibilities (resting and procession without reversion, procession and reversion without resting, and resting and reversion without procession) are equally unintelligible. Thus everything which is produced will both be distinct from its cause, while retaining some element of the cause within itself (or, as the

Neoplatonists preferred to put it in order not to subvert the proper metaphysical order of things, itself remaining partially in the cause), and seeking to revert towards its cause. These (non-temporal) sequences of procession and reversion are isomorphic mirror-images of each other, the one away from, the other towards, perfection (Props. 36–8).[38]
There follows a set of propositions (40–51) dealing with the 'self-constituted things' or *authupostata*. Proclus feels it necessary to posit such existents in order to secure self-sufficiency; and they must be intermediate between the Good (as Proclus here chooses to characterize the primary reality or One) and things which are simply caused (Prop. 40). The doctrine developed here is designed to carve out some elbow-room for human freedom, the need for which exercised Proclus just as much as his fellow Neoplatonists; it is probably owed to Iamblichus.[39]
Everything self-subsistent is self-constituted (Prop. 41), while all and only self-constituted entities are capable of reversion upon themselves (Props. 42–3). Such entities are without origin in time (Prop. 45), imperishable (Prop. 46), partless (Prop. 47), perpetual (Props. 48–9), and outside time at least as far as their existence is concerned, although not thereby necessarily excluded from temporal activity (Props. 50–1; the human soul,
end p.433

while being in a sense timeless, is still capable of temporal activity: Prop. 191). Proclus' arguments for the immortality of spiritual substances derive in part at least from Platonic sources (in particular *Phaedrus* 245c–246a: III.2*a*, **131**): generation requires pre-existing substance; and that substance cannot, on pain of regress, itself be subject to generation.
After a brief excursus on time and eternity (Props. 52–5) Proclus turns once again to consider the causal structure of reality (Props. 56–65). He first seeks to establish that in causal sequences (these sequences are metaphysical, and not temporal: contrast Philoponus, § *c* below) that

474 everything which is produced by the secondary things is also and to a greater degree produced by those prior and more causal things from which the secondary were produced. (56. 54. 4–6 Dodds)

Causes are prior to their effects, and their effects are greater in number than those of their consequents, since causes are definitionally of greater power (Prop. 57):

475 from this it is clear that Mind is also cause for all the things which Soul is a cause of, since it operates prior to it, and what Soul gives to the secondary things, Mind gives to a greater degree; and where Soul no longer operates, Mind illumines with its own gifts things to which Soul has not given of itself; for even the inanimate, in so far as it participates in form, participates in Mind and the activity of Mind. Moreover, the Good is also cause for whatever Mind is the cause of, but not vice versa: for even the privations of form derive from there, since everything derives from there [i.e. the Good]; but Mind, being Form, cannot give rise to privation. (57. 56. 8–16 Dodds)

These passages distil the essence of Proclus' system of reality. The Good produces Mind, which in turn produces Soul (and hence, by Prop. 56, the Good also produces Soul); but inanimate objects, although they are not produced by Soul, are produced by Mind, since they have structure, and that is a feature of Mind (equally, again by Prop. 56, they are produced by the Good). Finally, matter as such, having no structure, has no causal relation to Mind; but even it must owe its being to the Good, since everything does.[40] Succeeding propositions tease out some consequences of this doctrine, and that of the relationship between wholes and parts (Props. 66–74), not always with great lucidity. But we may disregard them and turn finally to that set of propositions (75–86) dealing directly with the relationship of cause to effect. Proclus begins by asserting that

> **476** everything properly called a cause transcends its effect; since if it were in it, it would either be a contributory part of the effect, or would require it somehow

> for its own existence, and would be in this way inferior to its effect. What exists in the result is a co-operative cause rather than a cause, being either a part of what is generated or a tool of the producer. . . . So every genuine cause, at least if it is more perfect than what comes from it and provides a measure for its coming to be, transcends the instruments, the elements, and in general everything which is called a co-operative cause. (75. 68. 30–70. 4 Dodds)

Elsewhere (e.g. *On the 'Parmenides'* 1059), Proclus distinguishes three 'genuine causes', the final, paradigmatic, and efficient, from two 'cooperative causes', the formal and material (cf. X.2*d*, **365**), sometimes adding the instrumental cause to the latter class (*On the 'Timaeus'* 1. 261: cf. X.1*b*, **354**; 3*b*, **368–9**; XI.3*a*, **401**; *b*, **402–3**; below, § *d*, **487–9**; on co-operative causes, cf. *Timaeus* 46c–e; III.4*b*, **144**). Proclus next argues (Prop. 76) that everything deriving from an immobile cause will be unchanging as regards its existence, since an immobile cause could not be a cause of variation; however, the products of causes which are themselves moving will be variable in this way. Thus unmoved movers will produce movers which move eternally; but the latter, *qua* causes, will themselves produce movement in other things variably. Proposition 77 offers an argument for PPA:

> **477** every potential existent advances to actuality as a result of something's being actually what it is potentially. . . . For the potential cannot itself drive itself forward into actuality, as it is imperfect; for if, being imperfect, it were still to be the cause of its own actual perfection, the cause would be less perfect than the effect which proceeds from it. Thus the potential as potential is not the cause of its own actualization. . . . So if the potential is to become actual, it must obtain its perfection from something else. (77. 72. 20–32 Dodds)

Proclus then argues by elimination that what it derives its actuality from must itself be actual, and actual in the same respect: thus PPA is exhibited as a form of PCS.[41] What is in fact required for causal activity is the concurrence of two sorts of potentiality, which Proclus proceeds to distinguish (Props. 78–9): the active power of the cause to bring about an effect, and the passive potentiality of the subject to receive it. This idea is not, in its essentials, un-Aristotelian; and a version of it may be found in the Stoics' discrimination of active and passive principles (VII.1*b*). Moreover, Proclus refers to the suitability of matter to receive form, a view we have also encountered elsewhere (§ *a* n. 31 above; cf. XI.2*b*, **397**). In fact

478 it is the nature of every body to be affected, and of every incorporeal to act. . . . Yet the incorporeal too is affected through association with the body, just as bodies too are capable of acting through their partnership with incorporeals. For every body in so far as it is a body is only divisible, and for this reason affectable . . . while the incorporeal, being simple, is unaffectable. . .

end p.435

Every agent possesses an active power, while body is in itself without quality or power, and hence will not act in virtue of being a body, but only in respect of the power for action within it; therefore it acts, when it acts, by participation in power. (80. 74. 27–76. 8 Dodds)

Bodies properly so-called are purely passive: when they appear to act they do so because they contain within themselves some portion of the incorporeal active principle which subsists in the physical body as a kind of irradiation, without thereby diminishing the original status of the incorporeal agent (Prop. 81). Thus the fundamental Neoplatonic metaphysical principles are turned to elucidate the relation between soul, body, and agency. Properly speaking, no body is ever an agent; it is merely the instrument for some further incorporeal power that happens to be invested in it. Thus Proclus turns Stoic theory on its head: for him, as for other Neoplatonists (and in a later age Berkeley), only immaterial realities are causes; although, unlike Berkeley, Proclus does not deny the existence of matter itself.[42]

In conclusion, it is worth briefly mentioning Proclus' attitude towards the dominant, Ptolemaic astronomical hypotheses of the day (cf. XI.1*a*). In his *Outline of Astronomical Hypotheses* (1. 1–6), he sets out to show that the pictures elaborated by Aristarchus, Hipparchus, Ptolemy, and others are as a matter of fact false of the actual mechanics of the heavens, although they may none the less function as useful (if flawed) heuristics (7. 50–8), a view repeated in his commentary *On the 'Timaeus'* (256b, 272b). Were the epicycles and eccentrics to have physical reality, they would defy the laws of his predominantly Aristotelian physics; and the proper, non-instrumental answer to the question why the heavens move as they do must begin from an understanding of their natures as ensouled bodies (284c; cf. XI.1*a*, **392**).[43]

(c) Philoponus

John Philoponus (*c*.490–*c*.570), a Christian Neoplatonist from Alexandria, was also deeply concerned with the proper explanation of celestial mechanics; and he was, if anything, more pessimistic than Proclus that a satisfying account was within the compass of human intellectual capacities. In *On the Creation of the World* (3. 4), he confesses his inability to explain why there should be the number of stars there are, or why they should be in their various places, or why they should be of different colours and magnitudes (a fact which he none the less ascribes to their being composed of different sorts of fire: 4. 12):

end p.436

479 but at least we all believe that God created everything beautifully and for a purpose, neither too much nor too little. In general we understand the causes of a few things only, and if people cannot tell us the natural causes for things evident, they should not inquire of us the causes of things which are hidden. (3. 4)

Philoponus does not doubt that there are final causes for the arrangement and structure of the heavenly bodies; but he does doubt that they are recoverable by us. Since this is so, we must rest content with whatever material and efficient accounts we are able to supply. In this regard, however, he was revolutionary. He sought to demolish the Aristotelian barrier between the celestial and sublunary realms, and in so doing affirmed that the heavens too, not being made of any special incorruptible element, were also subject to change and decay. Thus there is no obstacle to the assumption, in other respects an obvious one, that the heavenly bodies are fiery, and composed of combustible material. Philoponus rejects Aristotle's argument that, were the heavens composed of fire, fire would have long since overwhelmed the other elements (*Meteorology* 1. 3. 339^b16–40^a3), on the grounds that the same argument ought to apply to Aristotle's own fire-sphere itself (Philoponus, *Against Aristotle on the Eternity of the World*, [44] Fr. 52 Wildberg Simplicius, *On 'On the Heavens'* 80. 23–81. 11; cf. Fr. 54 W): the intensity of something's qualities does not necessarily increase proportionately to its mass (Frs. 53, 55 W). The heavenly bodies, then, are largely composed of fire, but also contain traces of the other elements in their purest forms (Fr. 56 W; cf. Plato, *Timaeus* 31b, 40a, 58c). Philoponus' contemporary and arch-rival Simplicius argued fiercely against these views and in favour of the world's eternity. When Philoponus maintained (on Platonic premisses: cf. *Timaeus* 31b) that it was reasonable to assume that the qualities of the heavenly bodies were not categorially distinct from those of the sublunary world, since the facts of colour and transparency at least seem the same in each realm (and hence argue for their fundamental material similarity: Fr. 59 W; cf. 57–8, 70–1 W), Simplicius responded with characteristic vehemence: such views are blasphemous and degrade the purity of the celestial realm (Frs. 59–60 W), while no empirical evidence has ever been observed which would suggest that the heavens are subject to decay (Fr. 57 W; Simplicius, *On 'On the Heavens'* 1. 3. 117. 21–118. 13; *On the 'Physics'* 8. 10. 1335. 2–16; cf. Aristotle, *On the Heavens* 1. 3. 270^b12–16). Philoponus, however, accepted that any decay was hard to detect: that fact is compatible with the heavens being extremely long-lived, yet ultimately destructible (Fr. 80 W).

end p.437

Aristotle had argued that the heavens could not be generable or destructible, since generation and destruction were from and into contraries, yet there is no contrary to the proper circular motion of the heavens (*On the Heavens* 1. 3. 270^a12–22). Philoponus agrees that it has no contrary in the sense of polar opposite; but it could surely have a privation, namely rest, and the only sense of 'contrary' in which it is unrestrictedly true that generation is from contraries is that of privation: there is no polar contrary, for example, to being a human being, yet that fact does not guarantee our immortality (Frs. 64–72 W; cf. Frs. 81–5 W).

As a good Christian, Philoponus subscribes to the view that the world was created literally out of nothing. That is, he rejects P1*b* (I.3*d*); things do not necessarily require pre-existing matter to be created from:

480 he allows, though, that in what he has said so far he has not refuted the view that it [i.e. the world] was generated from a pre-existing substrate; and so he wishes to show that the world came into existence out of what is not. And he refers in passing to a proof of this in his *Against Proclus*, although he presents the argument as an objection which he seeks to refute, saying: 'for if something were generated out of what absolutely is not, they say, it would follow that what is not exists, since it has changed into what is. But if', he says, 'someone says that the things which are generated out of nothing are generated from it in the same way as ships from wood, i.e. from what is not as a substrate for what is generated and which changes into it, then it will truly follow that what is not exists. But I do not suppose anyone to be so lacking in sense as to suppose that generation out of what is not occurs like this; rather each of the things generated comes into being in so far as it has not existed previously.' (Simplicius, *On 'On the Heavens'* 1. 3. 136. 14–26 Fr. 73 W)

Thus 'the non-existent' is not, absurdly, the material substrate for creation—rather creation *ex nihilo* simply is coming to be without a substrate: and Philoponus robustly asserts that, contrary to the prevailing ancient view, that is not an absurdity. When God generates from nothing, he generates both form and matter at the same time (Fr. 74 W); and Philoponus is right that such a supposition is not logically impossible. Again, a hallowed a priori causal dogma is rejected (compare Iamblichus: § *a* above).

Since there is no difference in type of substance between heaven and earth, Philoponus rejects Aristotle's view that the heavens move as they do in virtue of their composition from a distinct fifth element. He rebuts Aristotle's purely conceptual arguments for such a position in detail (Frs. 87–107 W); and while conceding that a fifth element was not conceptually impossible, it could neither be imperishable nor lack contraries as Aristotle had supposed, since the objects in the heavens are plainly differentiated (Philoponus, in Simplicius, *On the 'Physics'* 8. 10. 1330. 7–17);

end p.438

every physical body must be composite, and hence naturally subject to change and decay (1331–3 10–25).

Philoponus allows that there are such things as Aristotelian natural motions (cf. V.2*b*), but he argues that there is nothing absurd in the idea of bodies having more than one natural motion (*Against Aristotle*, Frs. 1–8 W), in the sense of distinct capacities for moving naturally under different circumstances. Thus fire, in addition to rising naturally to the periphery, may perfectly well naturally revolve there too (Frs. 9–17 W; cf. Xenarchus, X.4*a* n. 54); there is no need to suppose, as Simplicius and others did, that such movements are neither natural nor unnatural but supernatural. There is, however, a difficulty here: elsewhere (*On the 'Physics'* 3. 3. 378 25–8), Philoponus apparently suggests that the fire-sphere in its contact with the heavenly bodies moves not naturally but 'marvellously' (*huperphuōs*: literally, 'supernaturally'), which seems perilously close to designating it as supernatural (*huper phusin*) as Simplicius does (*On 'On the Heavens'* 35); and elsewhere its motion is compared to that of water in a revolving bucket

(Philoponus, *On the Eternity of the World against Proclus* 13. 5). I suggest, tentatively, that Philoponus' position is as follows: on *Aristotle's* view circular motion for fire should be natural, since it cannot be unnatural, and it must be one or the other. Philoponus believes that it is natural in one sense (the sort of motion which such objects are fit to receive), although not natural in the sense that other things being equal it will continue to move in that way without external force (Philoponus will elsewhere distinguish different senses and perspectives in which things may be seen as natural: see further below).

In any case, the natural motions are not determined by the prior existence of natural places; place as such possesses no power (compare Aristotle, *Physics* 4. 1., 208^b11): rather the bodies themselves have been endowed by the Creator with a disposition to seek their appropriate *relative* places, fire towards the periphery, earth around the centre, but not because the periphery happens to be located where it is (Philoponus, *On the 'Physics'* (*Corollary on Place*) 579. 19–82. 18[45]). Place, in fact, is nothing more or less than three-dimensional extension (567. 8–57. 9), having no essential properties of its own (cf. §2*a*, **439** above).[46]

Aristotle had held that motion in a void was impossible, since the speed of a moving body was inversely proportionate to the resistivity of the medium through which it was moving (*Physics* 4. 8. 215^a24–$^b22$; V.2*c*, **188**), and hence if the resistance of the medium were zero, the body would have

to move infinitely fast, i.e. instantaneously, which is impossible; and on this assumption, bodies of different weights would all move at the same (infinite) speed in the void, which is contrary to the hypothesis that their speed varies according to their weight (4. 8. 215^b22–216^a23). Finally, projectiles could not move in a void, since there would be nothing to maintain their motion once they had left the thrower's hand, and equally nothing to explain their ever coming to rest (4. 8. 215^a14–22; cf. V.2*c*, **189–90**).

Although he did not think that there was any actual void in nature (*On the 'Physics'* (*Corollary on Place*) 567. 32–3; (*Corollary on Void*), 675. 16–29), Philoponus rejected all of these contentions. First he denied that the motion of a body was *simply* inversely proportional to the resistivity of the medium: some of the moving object's speed was to be accounted for as being the result of its own internal nisus:

481 if the difference of motions arises not only because of the difference of the bodies through which the motion occurs, but also because of that in the movers themselves, if one is heavier and the other lighter, then clearly there will be a cause of the movers' unequal motion beyond the difference in motion that arises as a result of the difference in the bodies through which the motion occurs. But if this is the case, then clearly even if they were moving through a void their motion must be different too, since the motive cause within them is different. (ibid. 677. 13–21[47])

And a little further on:

482 if there are two causes of movers not all moving equally fast, the difference in what they pass through and the difference in the moving bodies, and the cause attributable to what they pass through is an obstructing cause (*empodistikē aitia*), then the cause ascribable to the difference in the moving bodies must be a productive cause (*poētikē aitia*). . . . So even if the obstruction is removed there nevertheless remains that cause of the unequal motion which is in the moving bodies, namely the productive cause. (678. 17–23)

Even in a void, the internal differences of weight between objects will still endow them with tendencies to move at different speeds: the extra cause of speed differential supplied by a resistive medium will *ex hypothesi* be absent, but that is no reason to suppose that all motions in a void will be infinitely rapid (cf. 680–1).[48]

So although Philoponus famously rejected Aristotle's claim that the speed of falling bodies was directly proportional to their weight (683), he did not suppose, as Galileo was to do, [49] that the speeds (or rather the

accelerations) of all bodies of whatever mass in free fall in a vacuum would be equal. His importance lies rather in his abandonment of the Aristotelian principle that 'everything that moves is moved by something' (*Physics* 8. 10. 266^{b}29; V.2*c*: **189**), [50] where that means not merely that it must have received some initial impulse (which Philoponus accepts), but rather that there must be some continuously acting external agent for all cases of motion, at least of forced or non-natural motion. Philoponus in any case rejects the simple dichotomy between natural and non-natural motions (and states of affairs). Things can be viewed as natural or otherwise only in relation to some particular frame of reference; and what may be unnatural in relation to one (disease or deformity for the individual) may yet turn out to be in accordance with nature as a whole (Philoponus, *On the 'Physics'* 2. 1. 201. 10–202. 16). Here Philoponus, following and developing a thought of Themistius (*On the 'Physics'* 2. 1. 37. 22–38. 18; cf. Simplicius, *On the 'Physics'* 2. 1. 271) to the effect that not every product of nature can be called natural in a strong sense, adopts the type of theodicy owed to the Stoics (VII.3).

In regard to projectile motion, Philoponus branches out in a striking new direction.[51] Philoponus denies that every moving object whose original cause of motion was external to it requires a persisting and continuously acting external force to perpetuate its movement. Rather the external mover is conceptualized as injecting a motive force into the object, a force which (at least if it is finite) gradually dissipates over time until the object ceases to move (Philoponus, *On the 'Physics'*, 4. 8. 639. 3–642. 26). Such motion is obviously perfectly possible in a void, as Aristotelian externally caused motion by contact clearly is not. This force is incorporeal, and directly impressed into the object (642); thus there is for Philoponus no Aristotelian differentiation between the types of body which are suitable for being motion-causers (*Physics* 8. 10. 267^{a}2–11: V.2*c*, **190**). Philoponus presses the explanatory advantages of his view against that of Aristotle:

483 if somebody throws a stone forcibly, is it by pushing the air behind the stone that he moves it forcibly in an unnatural direction? Or does the thrower impart some motive power (*kinētikē dunamis*) to the stone? If he does not impart any power to the stone but moves it simply by pushing the air, what need is there for the hand to touch the stone, or the bowstring the arrowshaft? (Philoponus, *On the 'Physics'* 4. 8. 641. 13–19)

No Aristotelian medium is required to operate as the continuing cause of motion; but neither is it needed to account for the fact that all such motions, even in a vacuum, must come to an end:
end p.441

484 just as you hold the thrust of the air responsible for unnatural motion, and say that such motion continues until the motive power which passes from the original pusher into the air is dissipated, I also hold that clearly if something is moved unnaturally in a void, then it too will move until the motive power imparted to it by the original pusher is exhausted. (644. 17–22)

This account of projectile motion has interesting consequences for the movement of the heavenly bodies: for they too, on Philoponus' account, must move because of a force initially impressed into them (*On the Creation of the World* 1. 12), a force, moreover, which cannot be infinite (Philoponus adapts to his own ends an argument of Aristotle's here: *Physics* 8. 10. 266^{a}24–^{b}27); thus even the heavenly bodies will eventually exhaust their impetus (Philoponus, in Simplicius, *On the 'Physics'* 8. 10. 1326. 38–1327. 35).[52] Their movements are natural in the sense that they do not require continuous external influence to keep them going. But they are not eternal; and their continuance requires a constant expenditure of the energy which has been stored inside them since the initial impetus was given.

Philoponus takes over from his Neoplatonist forebears some of the terminology they use to describe the relations between the various causal levels of reality (procession, rest, etc.: cf. Proclus, § *b* **473**, above). But for Philoponus they allude to temporal processes, in which force is communicated through time from mover to moved, rather than to a continuous functional relationship on the Procline model. Commenting on the Aristotelian doctrine (*Physics* 3. 3) that motion is in the thing moved rather than the mover, Philoponus writes:

485 the motive power which is impressed into the moved by the mover, by which I mean that force and activity (*energeia*) in respect of which something potentially movable is moved and brought to completion, is a single thing which starts the process from the mover, and has its end (*telos*), completion, and (as it were) its resting-place in the thing moved. For having come to be in what is moved potentially, it does not leave it, but, as it were, rests there and completes it, the completion being the production of force, namely motion. (Philoponus, *On the 'Physics'* 3. 3. 384. 33–385. 5)

There is a direct transfer of power from mover to moved, in contrast with Proclus' official Neoplatonist line, where the power strictly speaking remains in the cause (**476**).[53] Moreover, as a result of the transfer of power, the moved object becomes in a sense an autonomous mover: it is to its own internal conditions that we must now look for an explanation of its continued motion (here Philoponus parallels the views of the Stoics: VII. *d*–*g*),
end p.442

and it no longer owes anything directly to the influence of the originator of its motion.

This conceptualization of motion in turn allows Philoponus to develop a view of the genuine spontaneity of human action. The soul, for him, is an incorporeal force which permeates and animates those bodies which are fit to receive it (*On 'On the Soul'* 108–14, 329–32). It does not simply cause motion in response to certain desires and goals (as Aristotle had it: *On the Soul* 3. 10; cf. *Physics* 8. 6. 259^{b}1–10; see IV.4), but rather is capable (in some manner which is not made clear) of bringing those goals about themselves: the soul can directly choose both good and evil (Philoponus, *On 'On the Soul'* 18). Thus Philoponus sought to divorce the actions of the soul from direct causal dependence upon external influences, and thus to make it responsible in some strong sense for what it chose to do. This doctrine is of course in keeping with his Christian point of view; but it also reflects the extent to which, despite his obvious and avowed departures from it, Philoponus accepts a substantial part of his Neoplatonic philosophical inheritance.

(d) Simplicius

Simplicius too was a Neoplatonist, and, like Philoponus, a pupil of Ammonius, although he implies that their periods at Ammonius' feet did not coincide (*On 'On the Heavens'* 1. 2. 26. 19–21); he also studied with Damascius, another leading philosopher of the time and composer of Aristotelian commentary. But Simplicius was implacably opposed to Christianity, considering it a vulgar and intellectually disreputable religion; and he was one of the group of philosophers who fled into exile from Athens when the emperor Justinian closed the philosophical schools in 529 for religious reasons. Almost all of Simplicius' work, much of which survives, took the form, by this time the dominant philosophical genre, of commentaries on Aristotle, to whom he was a great deal more sympathetic than Philoponus had been. Thus he upheld the dominant Neoplatonic conception of an eternal world, arguing bitterly against Philoponus' attack *Against Aristotle on the Eternity of the World* in the course of his own commentary *On 'On the Heavens'*, as well as advocating an incorruptible fifth element and ascribing a natural power to place (Simplicius, *On the 'Physics'* (*Corollary on Place*) 623–36).
Simplicius is a commentator: and as such it is difficult (and not always profitable) to try to extract from his texts distinctive views of his own. I shall simply reproduce extracts from two passages having to do with our subject which are both representative of his style. The first comes from the beginning of his commentary *On the 'Physics'*:

486 and Plato posits three causes strictly speaking, the creator, the exemplar, and the end, and three auxiliary causes, matter, form, and instrument. However, Theophrastus, having given his account of the others, says: ‘after them came Plato, prior to them in reputation and ability, although later in time; and he was for the most part concerned with first philosophy, although he also dealt with the phenomena and offered an account of the natural world, in which he sought to make the principles two, the material substrate, which he called “all-receptive” [*Timaeus* 51a], and the cause and mover, which he ascribed to the power of God and the good’. Alexander, however, records that Plato spoke of three principles, matter, creator, and paradigm, even though Plato clearly posits the final cause when he says ‘let us then speak of the reason for which the composer composed generation and this totality: he was good, and for no good is there ever any jealousy of anything’ [*Timaeus* 29d–e]. And I think he clearly gave an account of enmattered form, where he spoke also of matterless Form . . . [here Simplicius quotes *Timaeus* 51e–52a: III.4*c*, **150**]; and thirdly he proposes matter in addition to these. (Simplicius, *On the ‘Physics’* 1. 2. 26. 5–25 Fr. 230 FHS&G)

The second passage occurs in Simplicius' discussion of Aristotle's exposition of the doctrine of the four causes in *Physics* 2. 3 (*On the ‘Physics’* 2. 3. 309–27; cf. IV.2*a*). The material cause is quickly dealt with (ibid. 309–10); but the formal cause requires more exposition. Simplicius (quoting a lengthy passage from Alexander) insists that Aristotle's enmattered form is not Plato's Form; but he also wishes, in good late Platonist fashion, to make Aristotle's picture compatible with Plato's. One must postulate above the mere existence of immanent enmattered forms a higher order of formal reality as ‘rational paradigms’ to explain the structure of the physical universe (312–14):

487 for this reason it is better to say that nature is a co-operative cause (*sunaition*), while the principal causes (*aitia prosechē*) of things subject to generation and decay are the motions of the heavenly bodies, in accordance with which they are directed. But higher are the psychical principles of these motions, and higher still the intelligible forms, out of which as primary is given to everything the illumination of the forms in accordance with the suitability[54] of the things which receive them. . . . It is clear that since the intellect, which, in accordance with the forms within it, assimilates the things which come to be to them, is the primary productive cause, then forms in the intellect will be the primary paradigms. (314. 9–23)

Separated paradigmatic causes must exist, since properly so-called causes must be actually distinct from their objects, while the ‘natural and generated form’ exists as a participation in matter; but, since participation is a relation of assimilation between participant and participated, there must be some independent source of participation for the enmattered forms

(317). Simplicius thus grafts the Platonic hypostasized Form onto the Aristotelian system with no sense of perceptible strain whatsoever, as indeed others had done before him (cf. X.2*d*, **365**).

Simplicius' exposition of efficient and final causes is orthodox and uninteresting (315–16): however, he takes issue with Alexander's account of the things which mediate between agent and goal (e.g. a doctor and health) as being efficient causes:

488 these things, namely the drug, the taking of the drug, purging, slimming, are both referred to the goal and come to be for its sake; thus the goal is a cause for their coming to be. They differ among themselves, in that some of them (exercise, purging) are actions, while others, such as the drug or the scalpel, are tools: but health is the final cause of all of them. Alexander calls them 'productive causes' of health; but it would be better to call them instrumental causes, for he himself agrees that the non-primary causes of motion are instrumental, and not properly productive. (316. 3–11)

Simplicius goes on to note that you might treat these intermediates as material causes, although he evidently prefers the instrumental designation (cf. Philoponus, *On the 'Physics'* 2. 3. 245–6). Thus Simplicius contrives to find the six Neoplatonic causes behind Aristotle's four. Commenting on Aristotle's remark (*Physics* 2. 3. 195^{a}3) that 'that [i.e. four], then, is pretty much the number of causes', he writes:

489 he adds the 'pretty much' either (*a*) because causes properly so-called are this many in number, although there are many incidental causes as well, as he goes on to say, or (*b*) as a result of caution, including the Platonic paradigmatic cause along with the principal causes, the productive and the final, and the instrumental with the co-operative causes, the material and the formal. But if the following is the correct arrangement of causes, such that the efficient and the final are primary by nature and properly causes, while the material and formal are co-operative causes, causes will reasonably be among the things said in many ways, but not among those distinguished as if from a single genus. (Simplicius, *On the 'Physics'* 2. 3. 316. 22–9)

Both Alexander (*On the 'Metaphysics'* 5. 2. 350) and Philoponus (*On the 'Physics'* 2. 3. 248) rightly interpret Aristotle's 'pretty much' in the manner of option (*a*) (cf. IV.2*a*). Alexander makes no mention of the paradigmatic cause, or of an instrumental cause as such at all; and while Philoponus argues that Aristotle omitted the paradigmatic cause only because he was here concerned with the world of change and decay, he also is reluctant to find an instrumental cause in Aristotle, preferring to think that Aristotle lumped the instruments together with the material (*On the 'Physics'* 2. 3. 241). But in Simplicius' syncretizing hands, Aristotle becomes an orthodox Neoplatonist.

As an interpreter of Aristotle, Simplicius is fair-minded and not unintelligent; but the example just given is typical of the way he

approaches Aristotle's texts. His work is massive and detailed, but rarely enlivened by any spark of intellectual brilliance or originality, or, indeed (apart from the polemical passages) by anything else. His texts are earnest, worthy, and largely tedious exercises in the frigid, baroque scholasticism into which Greek intellectual life had by this time largely declined.

4. Conclusions and Reflections

Plotinus' Neoplatonism had, obviously enough, intellectual roots in the various Platonisms of the preceding centuries, and took over much of their syncretism (§§1*a*–*b*).

But Plotinus' system was none the less an extraordinarily original achievement; and while its mystical tenor and its remoteness from the concerns of mundane physical explanation make it seem more of a precursor of the theologically dominated intellectual life of the succeeding Christian centuries than an heir to the tradition of Greek natural science which has been the principal focus of this book, his attempt to sketch a vision of reality in which everything is ultimately explicable in terms of its relationship with the supreme and indivisible One (§§2*a*–*d*) may be seen as the most extreme fruit of the general drive for explanation and intelligibility which informed the whole Greek enterprise, in all of its multifarious manifestations, from the earliest times onwards.

This concern to exhibit the whole of reality as a completely orderly, causally interrelated structure was taken over by his successors (§3*a*), and reaches its most formally developed form in the system of Proclus (§3*b*). Ironically, in view of the way in which things were to develop, it is the Christian Philoponus who shows the greatest interest in, and innovative flair for, natural explanation, taking on the still massive authority of Aristotle in questions of dynamics (§3*c*). And although Philoponus' trenchant criticisms of Aristotle were largely forgotten for centuries (although they were revived in the fourteenth century, and were a source of inspiration for Galileo), he is nevertheless a transitional figure: a man whose interests are in many ways those of his Greek patrimony, but whose iconoclastic views, in particular in regard to the eternity of the world, are informed by his desire to elevate the dogmas of the Christian faith above more secular philosophical reasonings.

Equally ironically, Simplicius, writing in exile from Athens, saw himself as the champion of the true theology and physics against the intellectually disreputable upstart Christian pretender, consoling himself with the belief that the Christians' triumph, and with it the fame of his *bête noire* Philoponus, would be short-lived, like the brief but deceptive blooms in the garden of Adonis (*On 'On the Heavens'* 1. 2. 25. 34–6). That comforting belief was, of course, utterly misplaced; it was the vigorous iconoclasm of

Philoponus which pointed the way forward. Even as Simplicius wrote, Christian hegemony—moral, political, and intellectual—was being established over the remains of the Graeco-Roman world, and with it came a preoccupation with the minutiae of theological doctrine, as well as a distaste, sometimes bordering on the pathological, for the speculations and investigations of worldly natural science. Greek learning did not die out immediately; commentaries on Aristotle continued to be written throughout the sixth century. But their authors were, increasingly, Christians, and increasingly concerned with Christian theological issues, while within a century or so the torch of scientific learning, and of learned commentary upon and elaboration of Greek science, had passed to the burgeoning, nascent world of Islam. The great intellectual adventure of pagan antiquity was over.

Appendix

List of Abbreviated Principles

This book makes use of a number of more or less formalized principles, which are labelled and rereferred to by abbreviations. This appendix seeks to smooth the reader's path through the text by reproducing here those abbreviations which are referred to elsewhere in the text, in the order in which they appear, with brief explanatory notes.

I. The Presocratics

PSR: Principle of Sufficient Reason.

The idea that nothing should occur without there being an explanation of why it occurs when and where it does; deployed negatively by Anaximander to account for the stability of the earth (I.2*e*); also invoked by Parmenides to refute the idea that things could have been created at a particular time (I.3*d*); and by Democritus (VI.1*e*).

PCS: Principle of Causal Synonymy.

The idea that an agent which produces some particular property must itself exhibit the property in question: if an agent *A* is causally responsible for some property *F* holding of *B*, then *A* must itself possess *F*, and make *B F* in virtue of its *F*-ness. First discovered in Alcmaeon's argument for the immortality of the soul (I.3*c*); so called after Aristotle (cf. IV.1*b*).

CP1: Nothing comes to be from nothing, a fundamental Greek axiom, enunciated by Parmenides (I.3*d*), but accepted by most Greek thinkers. It is ambiguous between

CP1*a*: nothing comes to be causelessly,

and

CP1*b*: nothing comes to be except from pre-existing matter,
an ambiguity exploited by the Eleatics, and diagnosed by Plato (III.1*b*: cf. P5 and P6 below) and Aristotle (IV.1*b*).

CP2: Nothing can be completely annihilated,
a corollary of CP1*b*, employed by Melissus (I.3*e*), but also generally accepted in the ancient world. The challenge for physicists now is to show how ordinary concepts of generation and destruction may be made compatible with CP1 and
end p.449

CP2, by analysing generation as fundamentally rearrangement, and insisting upon the appropriate conservation principles.

II. Science and Sophistry

CP3: For any *x*, any *y*, and any *F*, if *x* is the cause of *y*'s being *F*, then removal of *x* will contribute to the suppression of *F*,
a formalization of the Hippocratic allopathic principle that ‘opposites cure opposites’ (II.1*a*).

CP4: For any *x* and any *y*, if *x* is to interact with *y* then *x* and *y* must be of the same type, a formalization of Diogenes of Apollonia's view that only substances alike in kind can interact with one another (II.1*c*), later echoed in the view of the Stoics (and others) that only physical objects are capable of causal interaction (VII.1*b*; however, contrast the view of the Neoplatonists: XII.3*b*).

III. Plato

P1: The same effect cannot be produced by opposite causes,
a Platonic axiom (III.1*b*), which may be formalized as

P1*a*: if *x* brings it about that *y* is *F*, then it is impossible that there be some *z*, such that *z* is the opposite of *x*, and *z* can bring it about that *y* is *F*;
or, better,

P1*b*: if *x* brings it about that *y* is *F* in virtue of *x*'s being *G*, then it is impossible that there be a *z* such that *z* is *G**, and *z* can bring it about that *y* is *F* in virtue of its being *G**, where *G* and *G** are incompatible predicates; or, more simply,

P1*c*: if *G*-ness is responsible for *F*-ness, then *G**-ness cannot be responsible for *F*-ness.

P2: The same cause cannot produce opposite effects,
a corollary of P1, again expressible more formally as

P2*b*: if *x* brings it about that *y* is *F* (in virtue of its being *G*), then it is impossible that there be a *z* such that *x* brings it about that *z* is *F**;
and

P1*c*: if *G*-ness is responsible for *F*-ness, then *G**-ness cannot be responsible for *F*-ness.

P3: If *x* is responsible for *F*-ness, *x* cannot be *F**,
end p.450

i.e. a cause cannot bear properties incompatible with the properties it causes; which suggests

P4: If *x* is responsible for *F*-ness, *x* must itself be *F*,
i.e. PCS, which on certain Platonic assumptions entails (although it is not entailed by) P3.

A1: If x changes in respect of F, then there are times t_1, t_2 such that x is F at t_1 and x is not-F at t_2,
a formal expression of one temporal condition on property change: if something is to change property there must be a time when it has, and a (distinct) time when it lacks, that property. If change is to be continuous (i.e. if it is not to proceed by sudden jerks), and also to involve properties that can hold to different degrees, then A1 will become

A2: if x changes in respect of F, then there are times t_1, t_2 such that x is more F at t_1 and x is less F at t_2, and for any t_i, t_j, between t_1 and t_2, such that $t_i < t_j$, x is more F at t_i than at t_j;
i.e. at any stage of the process of losing a property, the object possesses that property to a greater extent than at any subsequent stage. These assumptions are explicitly endorsed by Aristotle (IV.1*b*, C).

P5: Nothing comes to be (F) from $_C$ not being (F),
and

P6: in every case of coming to be, something comes to be (F) from $_M$ not being (F).
P5 and P6 may both be true, in virtue of the different senses (causal and material) of the preposition 'from': i.e. opposites do not *cause* opposites (P5: compare CP1*a*); but coming to be requires the serial existence of opposite properties in some material (P6: compare CP1*b*); again an insight taken over and formalized by Aristotle.

P7: If x is a subtle cause for some property F, then x is invariably F.
P7 is Plato's account of a certain type of cause, that which, he thinks, properly explains its effect; and P7 is equivalent to

P8: If x is a subtle cause for some property F, then x is F by nature (i.e. essentially).
These principles are later exploited, negatively, by the Sceptics (VIII.3*c*).
The argument from design. The basic driving force behind Plato's cosmic teleology (III.4*a*; also accepted by Galen: XI.4):

(1) **[AD]** the world is such that it cannot have arisen except as a result of conscious creation;
(2) any such creator must be of supreme wisdom and goodness;

end p.451

hence
(3) every facet of that creation must exhibit the goodness of the creator's design.

IV. Aristotle: Explanation and Nature

C: x comes to be F from having been not-F:
Aristotle's basic characterization of the logical form of change (cf. A1 and A2: III.1*b*).

Two versions of a principle relating the notions of nature and frequency:
N: if q is natural, then given p, always or for the most part q;

N*: if it is natural for Fs to be Gs, then Fs are always, or for the most part, Gs.

VI. The Atomists

The issue between Stoic and Epicurean in regard to determinism and fixity of truth-value. Both hold the following conditionals to be true:
(1) if (*a*) truths are eternally fixed, then (*b*) if a proposition p is true at any time, p is true at all times;
and
(2) if (*a*) p is true at all times, then (*b*) p is necessary.
The Stoics argue from the truth of (1*a*) to the truth of (2*b*); the Epicureans, contrapositively, from the falsity of (2*b*) to the falsity of (1*a*). Carneades argued that neither inference was compelling, since the two following theses are distinct:

T1: each event is the wholly determinate product of prior causes (causal determinism); and

T2: if an event is going to happen it is already true that it will happen (logical determinism). And T2 does not entail T1.

VII. the Stoics

C1: x causes F.
This first attempt to capture the Stoic notion of cause is amended to

C1a: x causes y to be F: a corporeal agent x brings about in a corporeal patient y an incorporeal predicate F.
Stoic *aitia sunektika*, or containing causes, are sufficient for their effects:

C2: $(x)(y)$(if x causes y, then whenever x occurs, y occurs);
end p.452

but they are also coeval with them:

C3: $(x)(y)$(if x causes y, then whenever x is present, y is present);
moreover, the strength of cause and effect are tightly related:

C4: $(x)(y)$(if x causes y, then whenever x is present, y is present, and any variation in the strength of x is matched by an equivalent variation in the strength of y).

VIII. The Sceptics

The Sceptics' arguments against the reality of causes (and of signs) rest on a general theorem of relationality:

R: if *A* and *B* are correlative, then *A* cannot be apprehended before *B*.
Refutation of the sceptical argument involves the distinction between the following two causal theses:

CT1: Cause *C* is not apprehended prior to its effect;

and

CT2:*C* is not apprehended *qua* cause prior to its effect.
Even if true, CT2 does not entail CT1, which the argument requires. But in fact the most which is likely to be true is the trivial

CT2*: if *C* is the cause of *E*, then *C*'s being the cause of *E* is not apprehended prior to the apprehension of *E*'s being the effect of *C*.
The sceptical argument against causes (in particular antecedent causes: cf. Erasistratus, IX.1*c*, XI.2*b*) involves one or other of

CT3: if *x* is genuinely productive of *F*-ness, it should always produce *F*-ness in what it is in contact with;,
and

CT4: if *x* is genuinely productive of *F*-ness, it should be constantly *F*-producing.
Both theses are excessively strong (although CT4 is clearly stronger than CT3): and neither are justified.
XI. Science and Explanation
Galen's causal principles:

GP1: nothing occurs without a cause (cf. CP1*a*: I.3*d*);

GP2: everything comes to be from something which exists (CP1*b*);

GP3: nothing comes to be from the absolutely non-existent (which is a corollary to GP2, and a version of CP1*b*);
GP4: nothing is annihilated into the absolutely non-existent (CP2: I.3*e*);
end p.453

GP5: what is changed takes on a form similar to that which causes the change (cf. PCS: I.3*c*; P4: III.1*b*);

GP6: it is impossible that, when two bodies come together, they should not both act and be acted upon.

XII. the Neoplatonists

PPS: The Principle of Prior Simplicity.
This is the notion that, in a hierarchy of causal dependency, the higher the element, the more simple it will be. This requires a restriction on PCS (for the property of complexity), but it is linked with

PPA: the Principle of Prior Actuality.

The idea that if something is causally responsible for something else, it must be prior to it in actual existence. This owes a great deal to Aristotle, and involves a version of PCS.

LMT: The Law of Mean Terms: between any two completely distinct levels of reality there must be some intermediate stage sharing in some of the characteristics of each of them. LMT is central to Proclus' development of the Neoplatonic picture of reality.
end p.454

Archiv für Geschichte der Philosophie

Abbreviations

AGP	
AJP	*American Journal of Philology*
ANRW	*Aufstieg und Niedergang der römischen Welt*
BACAP	*Proceedings of the Boston Area Colloquium in Ancient Philosophy*
BHM	*Bulletin of the History of Medicine*
BICS	*Bulletin of the Institute of Classical Studies*
CAG	*Commentaria in Aristotelem Graeca*
CJP	*Canadian Journal of Philosophy*
CMG	*Corpus Medicorum Graecorum*
CP	*Classical Philology*
CQ	*Classical Quarterly*
HSCP	*Harvard Studies in Classical Philology*
JHS	*Journal of Hellenic Studies*
JP	*Journal of Philosophy*
OSAP	*Oxford Studies in Ancient Philosophy*
PAS	*Proceedings of the Aristotelian Society*
PCPS	*Proceedings of the Cambridge Philological Society*
Phron.	*Phronesis*
PQ	*Philosophical Quarterly*
QJRAS	*Quarterly Journal of the Royal Astronomical Society*
RM	*Review of Metaphysics*
RUSCH	*Rutgers University Studies in Classical Humanities*

END