2

Ptolemy's Mathematical Models and their Meaning

Alexander Jones

Department of Classics, University of Toronto

In the middle decades of the second century of our era, a Greek-speaking Egyptian living in the town of Canopus close to Alexandria carried out a massive scientific program centring on the writing of about a dozen books on astronomy, astrology, optics, harmonics, and cartography.[1] Unlike his near contemporary Galen, Ptolemy evidently did not lead the sort of career, and certainly did not have the self-trumpeting personality, that would procure notoriety among one's contemporaries, and so we know scarcely anything about his life. But his works were well enough appreciated, in spite of their severe style and uncompromising technicality, so that the great part of them were preserved, almost the sole remnants of their kind of scientific writing from antiquity. Though ranging widely in subject matter, these books revolve around two great themes: mathematical modelling of phenomena, and methods of visual representation of physical reality. In the following, I wish to consider what Ptolemy thought the relationship was between his models and the physical nature that he was describing.

To begin, let us look briefly at what his predecessors made of this question. The explanations of phenomena offered by Greek physical science varied greatly, but they were often framed in terms of two broad principles: first, that change in matter can be reduced to the operations of a small number of fundamental qualities, typically hot, cold, wet, and dry; and secondly, that the phenomena can be modelled by mathematical objects. These principles tended to be regarded as mutually exclusive, so that certain phenomena were referred to qualitative and others to quantitative explanation.

What decided which kind of rationale was appropriate in any particular situation? The historical reality was surely that people stuck to whichever

[1]For the biographical data on Ptolemy see Toomer 1987. (Given the informal nature of the present paper, I have thought it appropriate to furnish the text with references only to translations of the pertinent works and to a few particularly helpful works of modern scholarship. The translations of passages quoted in this paper are my own.)

kind seemed to work best for the subject matter. Certain areas of experience lent themselves more obviously to mathematical modelling than others; in particular, the motions of the heavenly bodies, the esthetics of musical intervals, and the visual perception of shape, distance, and motion were sensed to have a quantitative regularity not shared by physical change in materials such as heating, melting, or burning, which on the contrary seemed to have a fairly direct connection with transference of the qualities hot, cold, wet, and dry.

Aristotle's cosmology, with its inner globe of more or less stratified earth, water, air, and fire enclosed in an outer spherical shell of ether, was in part motivated by this polarity, and in return gave it an *a priori* rationalization. The matter of the heavens–the part of the cosmos where the stars, planets, sun, and moon dwell–is of a different kind from the four mundane elements, subject to a different natural motion (circular revolution as opposed to motion towards the cosmic centre), and not subject to any other kind of change. Aristotle's ether has the power to force change in other things, but considered by itself, its only property is eternally regular circular motion. Hence an Aristotelian astronomy has everything to do with mathematics, and nothing to do with elementary qualities. Earth, air, fire, and water, on the other hand, can be forced by an external agent to move in any direction or to change properties, and, moreover, these processes vary unpredictably in degree and duration. This is why, even if the continual changes among the four elements–including life itself–can be traced back through a chain of cause and effect to the physical action of the heavenly bodies (most importantly the sun's annual revolution alternating between north and south), terrestrial phenomena are not as regular and periodic as the celestial revolutions:

> We see that when the sun comes closer, coming-into-being takes place, and when it recedes, ceasing-to-be takes place, and each happens in equal time. . . . But it often happens that things cease to be in a shorter time because of the mixture of things with one another; for since their matter is not uniform and not the same everywhere, necessarily their comings into being too are not uniform, and some are faster and some slower. . . . (Aristotle, *De Gen. et Corr.* 336b16)

Aristotle's cosmology thus explains why we can have a mathematical astronomy. It does not, however, account for the possibility of mathematical sciences dealing with special aspects of the world of the four elements, although Aristotle recognized that possibility, since he classified optics and harmonics, along with astronomy, as sciences embedding mathematics, or indeed as *branches* of mathematics (*Physics* 194a6). Here and there in the Aristotelian corpus we encounter *obiter dicta* confirming that Aristotle recognized that mundane phenomena could be subject to mathematical constraints, for example in the following passage where he speculates on a possible analogy between harmonic theory and colour theory:

> We have to discuss the other colours [besides white and black], distinguishing the number of ways that they can arise. Now white and black can be placed side by side in such a way that each one cannot be seen because of its tiny size, but the product of the two becomes visible in this way. This cannot appear either as white or as black. But since it must have *some* colour, and it cannot be either of these, it must be a mixture and some different form of colour. In this way one can suppose that there are more colours besides white and black, and that they are numerous in accordance with ratio. For they can lie next to each other in the ratio three to two, and three to four, and in ratios of other numbers; and others can be wholly in no ratio, but incommensurable by some excess and defect. And it is possible that these things subsist in the same manner as (musical) concords; for the colours that are in numbers that form good ratios, just like concords in the other context, would seem to be the most pleasant of colours, for example sea-purple, red, and a few others like these, for the same reason that there are just a few concords, while those that are not in such numbers are the other colours. (*De Sensu* 439b19)

But an analogy is not an explanation, and we are left in the dark as to why simple ratios of whole numbers should have a special status in a world of geometrically continuous matter and change. Similarly one is left wondering why vision follows straight lines if it is in fact a process of continuous change in nonuniform matter.

For a working scientist of the Hellenistic or Roman periods in search of a broad rationalizing framework in which to set his own theorizing, Aristotle's cosmology and conception of matter were not the only ones on offer. In the first place, Epicurus revivified atomism into an elaborate, strictly materialistic physics in which all matter and change are reduced to the chance motions of eternal atoms, endowed with a minimum of properties (shape and size), in an infinite void. Epicurus has sometimes been portrayed as a prophet of science; in reality he was no friend to the sciences of his time. He endeavoured to show how the phenomena for which the astronomers sought unique explanations could result from numerous different physical situations, any of which might be temporarily valid at some time and place within his boundless universe; his theory that vision occurs by means of films of atoms that continually peel off bodies and fly off in all directions would not have stood up long to the scrutiny of a practitioner of geometrical optics; and in general he contemned any inquiry into nature that was not subordinated to his ethical goals, freeing humanity from avoidable pain and fear.

The physics of the Stoics was closer to Aristotle's. We find again an insistence that matter is geometrically continuous and reducible to variable mixtures of a restricted number of fundamental stuffs, which at one level of analysis prove to be the familiar earth, air, fire, and water. The Stoic cosmos is finite and spherical, but there is no outer shell of special unchanging matter for the heavenly bodies, and the cosmos in its present differentiated state has a finite span of life. On the other hand, although Stoic physics is strictly materialist, its cosmos is orderly and rational. The organization of the cosmos and its parts is effected by pneuma, a vital mixture of fire and air that extends in varying degrees throughout the cosmos and that has the power to "tense" the bodies with which it is intermixed. In place of a reductionist explanation of the mathematical behaviour of phenomena, we encounter a deistic appeal to the will of the cosmic mind. (It should, however, be kept in mind that our sources for Stoic physics are less satisfactory and more controversial than those for Epicurean physics, and in any case Stoicism was considerably more open to innovation than Epicureanism with its *ipse dixit* deference to its founder's pronouncements.)

Alongside these more or less coherent systems there existed a looser tradition of physical speculation, which we call "Peripatetic" because its most prominent known advocates, in particular Theophrastus and Strato, were close associates and followers of Aristotle. This was an eclectic approach, grounded in observation and analogy, and again materialistic. Properties of matter and processes of change are explained in fairly mechanical terms, for example by supposing that materials can be composed of particles that can be packed loosely or tightly, but the particles lack the permanence of true atoms and are less denuded of innate characteristics. Aristotle's fifth element seems to have won no following; the heavens were instead supposed to be composed mostly or entirely of fire. This fire might be endowed with special properties, perhaps, but the divide between the celestial and mundane spheres was inevitably blurred.[2]

Such were the main lines of physical thought evolving during the century following Aristotle's death. It was also at this time that the earliest surviving works that treat physical problems using mathematical models were written. These include works on astronomy by Autolycus, Euclid, and Aristarchus, works on statics by Archimedes, and works on optics and harmonics by (or at least ascribed to) Euclid. What is striking about these works is not only the attempt to deduce phenomena through explicit axiom and theorem structures, but also the fact that these works seem deliberately to evade physical interpretation of the axioms.

One would dearly like to know what developments the subsequent three and a half centuries brought. The state of evidence is far from encouraging. Thus, of the numerous books written by undoubtedly the most important mathematical scientist of this period, Hipparchus, we possess only one, and with scarce gratitude and less justice we tend to dismiss that work as atypical and uninteresting. Among the philosophers, Posidonius stands out as a writer who undoubtedly exerted a considerable influence on physical thought. One recognizes in some of the reports of his lost writings the tincture of Peripatetic

[2]The rejection of Aristotle's fifth element is ably discussed by Falcon 2001, 121-183.

physics in his Stoicism for which he was later criticized; it is harder to discern a serious engagement with the mathematical sciences.

For that, we must turn to Theon of Smyrna, a Platonist philosopher of far lesser distinction than Posidonius, but one with the accidental merit that a large part of one of his books, *The Mathematics Useful for Reading Plato*, has come down to us.[3] Theon's mathematics embraces harmonics and astronomy, and the long astronomical section is of particular interest here. Theon exposes, with geometrical demonstrations, the epicyclic and eccentric models as assemblages of circular paths in the plane; but he insists that these circles are not mere abstract conceptions but stand for spheres of ether such that, for example, an epicycle is a rotating sphere nested in the gap between two concentric spherical shells which revolve together, bearing the epicycle with them. Theon was a mere generation older than Ptolemy, but this is enough to establish that the revival of Aristotle's etherial spheres and their adaptation to non-homocentric models was not due to Ptolemy, though it may have been fairly new science in his time.

It makes sense in several ways to begin considering Ptolemy's attitude to mathematical models in the context of his astronomy. This was the science closest to his heart, the only one on which he is known to have written a multiplicity of books. His central treatise on astronomical modelling, known to us as the *Almagest*, preceded most of the others, yet it followed upon a quarter-century of personal observation and analysis.[4] It is also a monumental piece of reasoning, much more complex and at the same time more structurally unified than his other large works.

The models with which the *Almagest* is concerned are kinematic geometrical constructions built up from circular motions representing the paths travelled by the heavenly bodies (the sun, moon, planets, and fixed stars). Most of the bulk of the *Almagest*, and most of its mathematics (in the usual sense of the word), is devoted to determining the radii, rotational velocities, and orientations of the components of each model. These parts, taken in isolation, leave open the question whether the circles in the diagrams stand for some sort of physical bodies in motion, or whether they are just abstract analytical components of a complex motion which the heavenly bodies perform due to undetermined physical causes.[5] We can at least dismiss a third option, that they are mere computational devices with no necessary relation to what the heavenly bodies really do, but by which one can reproduce the phenomena seen by a terrestrial observer; Ptolemy's treatment of parallax and eclipses depends on the assumption that his lunar and solar models correctly describe the distances of the sun and moon from the earth as well as their directions from the observer.

---

[3] The most reliable translation is Dupuis 1892.

[4] Toomer 1984.

[5] On the question of Ptolemy's realism in the *Almagest* and *Planetary Hypotheses* see Lloyd 1978.

However, the broader deductive structure of the *Almagest* decisively commits Ptolemy to believing that his circles stand somehow for material bodies, even if it is not made explicit precisely how they do so. This may be seen by examining how Ptolemy arrives at each model *before* turning to the deduction of its numerical parameters. I take as an illustration Ptolemy's model for the moon, since Ptolemy's presentation of this model in Books 4 and 5 of the *Almagest* is particularly explicit about the stages by which the model is worked out.

Ptolemy starts out in *Almagest* 4.5 with a working hypothesis, which he warns us will later be disproved, that the moon has a "single and invariant" anomaly, that is, that its apparent progress along the ecliptic has a periodic variation that always repeats exactly. He asserts that two models identically produce this phenomenon. In one model (Fig. 1), the centre of an epicycle E travels eastward along a circular deferent concentric with the earth T with uniform angular velocity (relative to an arbitrary stationary radius from the earth's centre), while the moon M travels along the epicycle uniformly in the opposite direction (relative to the radius from the earth's centre to the epicycle's centre). The angular velocity of the moon on its epicycle is slightly less than that of the epicycle's centre on the deferent. In the other model (Fig. 2), the centre of an eccentric circular orbit C revolves with a slow uniform westward motion along a circle concentric with the earth T, while the moon M travels along the orbit with a uniform eastward motion (relative to the radius from the earth's centre to the eccentre's centre). The two models, as Ptolemy proves, are kinematically interchangeable; that is, any set of positions in space of the moon for specific dates generated by the one model can be generated identically by the other. Moreover, Ptolemy knows already that they are both incorrect, because the lunar anomaly is not simply periodic. Nevertheless, Ptolemy selects the epicyclic model as the basis for a preliminary lunar theory in which the numerical parameters are determined by analysis of observations of lunar eclipses. It is noteworthy that Ptolemy makes a point of showing that the same parameters result from several different sets of observation reports, thus establishing that the preliminary model is computationally valid for all eclipses (and by extension, all oppositions).

The motivation for Ptolemy's selection of the epicyclic model only becomes fully evident when he shows in *Almagest* 5.2 how it is defective. He finds that the "equation," or difference, between the moon's observed position and its mean position (that is, the direction to the epicycle's centre according to the model) is in general greater than the model predicts, with the discrepancy vanishing when the moon is at 0° or 180° elongation from the sun and maximum when it is at ±90° elongation. In an epicyclic model the equation is explained by the planet's motion on the epicycle, so that the new phenomenon (essentially equivalent to "evection" in later lunar theory) would amount

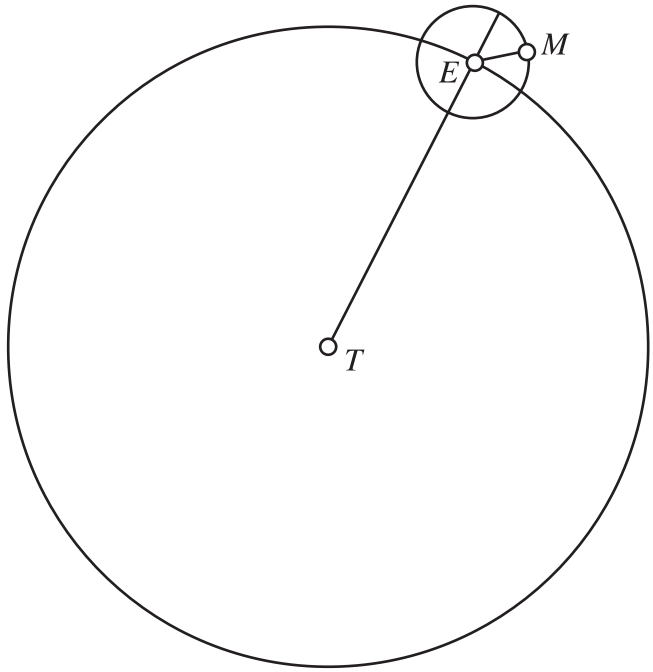

Fig. 2.1. Ptolemy's simple epicyclic model for the moon.

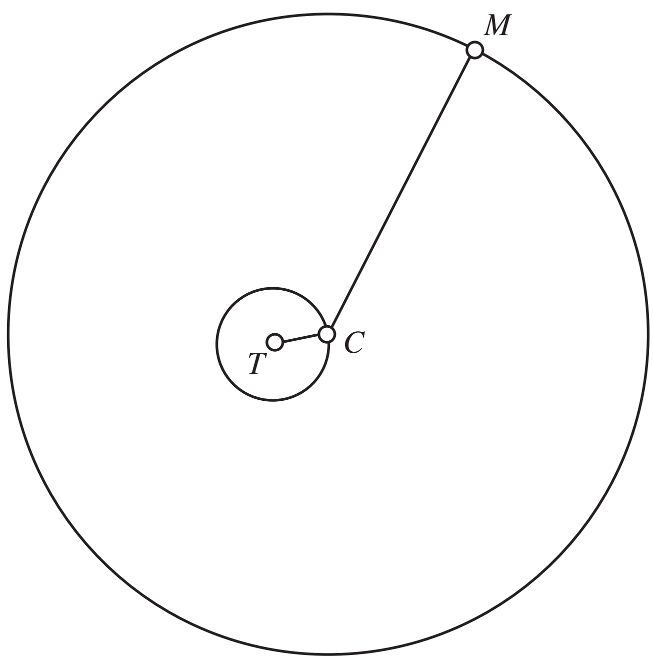

Fig. 2.2. Eccentre model equivalent to the model of Fig. 1.

to an apparent enlargement of the epicycle.[6] Ptolemy accounts for this easily (Fig. 3) by replacing the concentric deferent in the preliminary model with an eccentric deferent, the centre C of which revolves around the earth T at a rate such that the epicycle's centre E (which is still revolving uniformly as seen from T) is furthest from the earth whenever the mean moon and the mean sun are aligned or diametrically opposite. Now it is true that if Ptolemy had employed an eccentric orbit to effect the anomaly in Book 4, he could have corrected the model in Book 5 by adding an epicycle (or even a second, independent eccentricity), but the relation of the components to the phenomena would have been much less intuitive. And in any case Ptolemy fine-tunes the model, for closer agreement with observations, by stipulating that the moon's motion on its epicycle is uniform as measured relative to a revolving radius, not drawn from the centre of the deferent C or from T, but from a distinct point D such that T is always the midpoint of C and D. This could not, I believe, be translated in any straightforward way into a model in which the primary anomaly is produced by an eccentre.

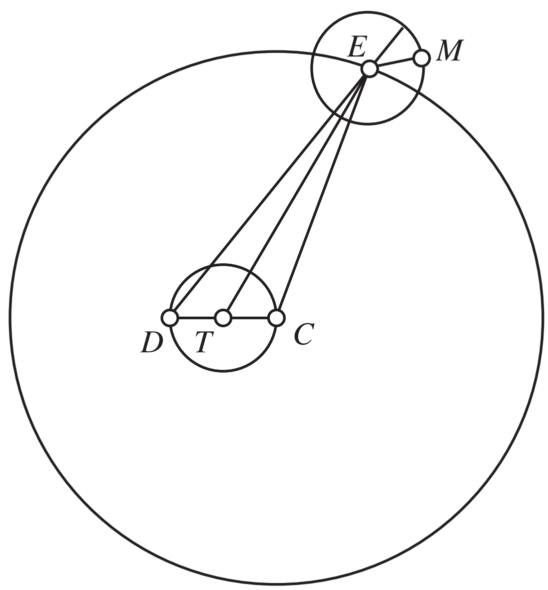

Fig. 2.3. Ptolemy's eccentre-and-epicycle model for the moon.

Thus all the stages from the selection of a basic model type to the final model are motivated in Ptolemy's exposition by the requirements of agreement with observations, simplicity, and a clear one-to-one correspondence of the

[6]For the relationship between Ptolemy's so-called second anomaly of the moon and the component called "evection" in modern lunar theory, see Neugebauer 1975 v. 3, 1108–1109.

elements of the model to the basic facts about the moon's motion. A similar account could be given for Ptolemy's deduction of the models for the sun and the five planets. Ptolemy makes no appeal in these parts of the *Almagest* to physical constraints arising from the corporeal nature of the model. But this is because those constraints have already been taken into account at a still earlier stage, the decision to build all the models out of uniform circular motions, which is made once and for all in *Almagest* 3.3, just before the first discussion of the sun's model. Here Ptolemy writes:

> The next task being to exhibit also the apparent anomaly of the sun, the assumption must first be made that the shiftings of the planets [including the sun and moon] in the trailing direction of the heavens [i.e., westward] are uniform, just like the movement of the whole [heavens] in the leading direction [i.e., the daily eastward rotation of the heavens], and they are circular by nature, that is, the straight lines that are imagined as leading the heavenly bodies or their circles in their revolutions sweep out in all cases equal angles in equal times with respect to the centres of each one's revolutions, while the apparent anomalies pertaining to them are produced by the positions and arrangements of the circles *on their spheres*, by means of which they make their motions, and nothing in nature really occurs that is foreign to their *eternity* in connection with the imagined irregularity of the phenomena.

This is one of only a handful of references in the *Almagest* to the circles in the models as being on the surfaces of spheres; when he does this, it is always in a matter-of-fact way, implying that the reader will already be familiar with the conception. In this particular passage Ptolemy uses language connecting the idea of uniform circular motion with physical nature and eternity, so that ether, though not explicitly named, is inevitably called to mind.

And this brings us back to Ptolemy's very first chapter, *Almagest* 1.1. Here he defines the science of which his subject matter is a part, which he calls "mathematics" (the *Almagest*'s real title is *Mathematical Composition*), as the study of shapes and spatial movements in all kinds of bodies, whether eternal and etherial or perpetually changing and composed of the four elements. Mathematics offers "sure and unshakeable knowledge," and when concerned with the etherial heavens, this knowledge is as eternal as its objects. In other words, the conviction that the heavens are composed of etherial bodies, which are by their composition both eternal and subject to no kind of change except circular revolution, guarantees the legitimacy and truth of the kind of reasoning that the *Almagest* embodies. It is noteworthy that, while practically every other theoretical hypothesis in the *Almagest* is justified by some empirical argument, the hypothesis of the etherial nature of the heavens is given axiomatically at the beginning.

His claim to be arriving at "sure and unshakeable knowledge" in the *Almagest* turns out in practice to have certain limitations. Numerical parameters are, by his confession, knowable only approximately, and in particular the rates of rotation become more precisely known as we accumulate a longer temporal span of observation reports. Ptolemy does not say outright whether he believes that his specific model structures (exclusive of their numerical parameters) are certainly valid. What the *Almagest* does affirm through its broad plan is that Ptolemy's models suffice to explain all the known phenomena of the heavenly bodies, including eclipses, planetary retrogradations, and visibility conditions. But the impression of finality is moderated, not only by the way that Ptolemy recounts his discovery of the moon's evection (might there be other phenomena waiting to be noticed?), but also by his passing reference to alterations he had only lately made in his models for Mercury and Saturn. In a famous passage towards the end of the work (*Almagest* 13.2), he justifies the complexity of his models for the latitudinal motion of the planets by affirming that the principle of simplicity in models should not be allowed to override the necessity to account for the phenomena, since what seems complex to us with our experience of the imperfections of mechanisms built from the four mundane elements may be simple to essences that are eternal and free from hindrance. Implicit in this is his confidence that his models really are the simplest that can be brought into agreement with observation.

Ptolemy's reticence regarding any but the most fundamental properties of ether and regarding the way in which the geometrical objects that constitute the *Almagest* models are supposed to be instantiated in etherial "spheres" in the actual heavens may be due partly to a reluctance to digress from the core subject matter of the book, but another reason may be that he had not yet given these topics much thought (just as he tells us in *Almagest* 2.13 that he has not yet worked out the list of geographical locations that he eventually delivered in the *Geography*). In a much later work, the *Planetary Hypotheses*, Ptolemy has considerably more to say about the spheres.[7]

The *Planetary Hypotheses* is ostensibly an exposition of the *Almagest* models, with some revisions, described in a manner that will be helpful for people who wish to make demonstration models or planetaria, with the parts either manually adjustable to their positions at any date or driven by a mechanism. After a first book in which Ptolemy sets out the parameters of all the models individually and proposes a scheme for nesting the models one inside the next, from the moon's model outwards to those of Saturn and finally the fixed stars, he turns in Book 2 to a consideration of the models as three-dimensional corporeal objects, that is, the "spheres" alluded to in the *Almagest*. Here Ptolemy engages in an extended discussion of his notion of how etherial matter works.

[7]The original Greek text of the *Planetary Hypotheses* is extant only for the first part of Book 1, for which see Heiberg 1907, 70–106; there is no reliable modern translation from the Greek. The whole of Book 1 in the medieval Arabic translation is edited and translated into French in Morelon 1993. For the Arabic text of Book 2, one currently depends on the German translation of L. Nix in Heiberg 1907, 111–145 and the facsimile of a manuscript in Goldstein 1967. Murschel 1995 is an excellent synopsis of the work.

Etherial bodies, he says, are subject to no external influence or alteration. To each independent motion in the kinematic models there corresponds a rotating etherial body incited into motion by the power of the visible heavenly body that it bears. These visible bodies (i.e., the sun, moon, planets, and stars) are the same in composition as the matter that surrounds them. They differ, however, in that they issue rays that have a power to penetrate other bodies, analogous to our intellects and vision. Moreover, their ability to set their spheres in motion is analogous to the power of our minds to cause our bodies to move; but in the celestial case the movement is utterly effortless.

Aristotle's cosmology had been strongly influenced by Eudoxus' astronomical models, in which the motions of the heavenly bodies were produced by combinations of circular motions all concentric with the centre of the cosmos; hence he could ascribe to ether a "natural motion" always perpendicular to any radius from the cosmic centre (in contrast to the natural motion of the four elements, which is always rectilinearly centripetal). For Ptolemy this cannot do, but he proposes a novel principle, that what Aristotle had characterized as natural rectilinear motion is in fact only natural to a body that has been removed from its "natural place." The etherial bodies, being already in their natural place, are subject to no tendency to migrate up or down, but are free to stand still or rotate effortlessly. "Mathematics" (i.e., deductive mathematical astronomy in the style of the *Almagest*) allows for two possibilities for the shapes of the etherial bodies. On the one hand, they can be spherical shells and solid spheres; but if so, they do not have to be imagined as being driven in their rotations by their axes, as in a mundane machine. Indeed, the entire polar regions of the spheres seem to Ptolemy to be superfluous to their motions, so that he is prepared to restrict the mobile bodies to equatorial slices of spheres and spherical shells, so-called "tambourines" and "rings," which are presumably sandwiched between regions of stationary ether. Since distinct etherial bodies can slide freely against each other, there is no need to imagine "unwinding" spheres that cancel out the revolutions of outer spheres, such as Aristotle had imposed on his mechanistic interpretation of Eudoxus in *Metaphysics Λ*.

When we come to the detailed description of each heavenly body's physical model, we find that the basic conception is similar to Theon of Smyrna's, but extended to include eccentric as well as epicyclic motions. Fig. 4 (a cross-section through the plane of the moon's orbit) shows how Ptolemy conceives of the arrangement of etherial bodies that bring about the moon's motion, on the assumption that the bodies are complete spheres or spherical shells. The entire apparatus must be thought of as being spun about the poles of the celestial equator with the daily rotation of the heavens. The outermost shell A rotates around the poles of the ecliptic with the slow motion of the moon's nodes. Within this is a shell B rotating around the poles of the inclined plane of the moon's orbit at the rate that, in the model of *Almagest* Book 5, the centre of the moon's eccentre revolves around the earth relative to the nodes. Cut out of shell B (and actually dividing it into two noncontiguous parts) is an eccentric shell C that has embedded within it the solid epicyclic sphere D. C and D together revolve uniformly as seen from the centre of the cosmos T with the rate that the centre of the epicycle revolves around the earth in *Almagest* 5. Finally, the epicyclic sphere rotates, carrying embedded close to its surface the moon M itself, producing the primary anomaly. This physical model is wholly consistent with the *Almagest* model, except that Ptolemy abandons the special radius with respect to which the moon's regular revolution on the epicycle is reckoned, instead stipulating that the moon's revolution is uniform relative to the radius from the centre of the cosmos. At the beginning of the *Planetary Hypotheses*, Ptolemy writes that the models as set out in this work incorporate some revisions to the *Almagest* models based on newer analysis of observations, but also that he is making some minor simplifications purely for the sake of an easier construction of demonstration models; one is left uncertain which kind of change is being made here in the lunar model.

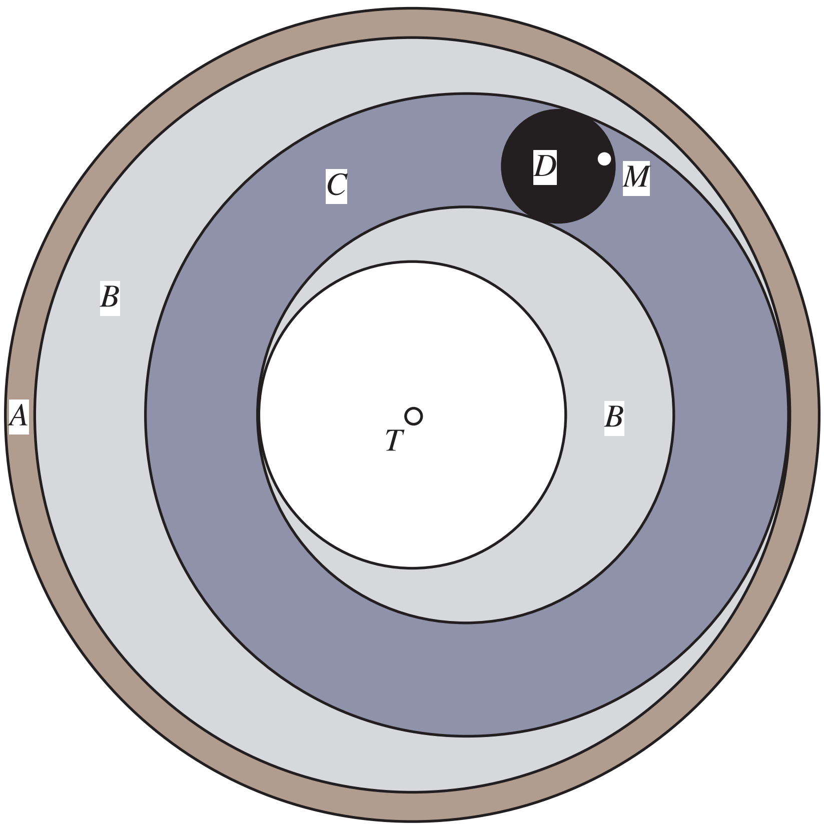

Fig. 2.4. Cross-section of Ptolemy's etherial-spheres model for the moon.

To the extent that the *Planetary Hypotheses* is intended as a description of the reality of the heavens (as opposed to its professed purpose of giving designs for didactic three-dimensional illustrations that we can construct), the pronouncements in the book are more equivocal than those of the *Almagest*. Ptolemy is quite sure of the etherial composition of the heavens, and also quite sure of the fundamental geometrical structures of the celestial motions; but the specific way that these geometrical structures are embodied in the etherial matter is open to alternatives (complete spheres or equatorial slices, spinning driven by cosmic souls or by planetary rays or by the axes). Ptolemy generally tells us which way he is inclined to choose, but these topics are not the province of the "unshakeable knowledge" of mathematics.

With the *Tetrabiblos* (a work written after the *Almagest* but probably well before the *Planetary Hypotheses*), Ptolemy turns from pure contemplation of the celestial realm of ether to an investigation of the action of the heavens upon the world of the four elements.[8] The fundamental assumption, comparable in its role in Ptolemy's astrology to the hypothesis of the uniform circular motion of ether in his astronomy, is propounded in *Tetrabiblos* 1.2:

> The fact would appear utterly obvious to everyone through even a few considerations that some power is given forth and reaches from the etherial and eternal nature to all the region around the earth, which is in all respects subject to change, with the first elements below the moon, i.e., fire and air, surrounded and directed by the movements in the ether and surrounding and directing all the rest, i.e., earth and water and the plants and animals within them.

There are, however, important differences between these fundamental hypotheses. In both, the etherial matter is simply a given. But the property of uniform circular motion in the *Almagest* is justified on *a priori* grounds (circular rotation being the only kind of eternally unchanging motion that can be conceived), whereas the property of exerting a power of change on the four elements is argued directly from empirical facts. Ptolemy backs it up with a series of examples of situations where laymen know perfectly well, and act on the knowledge, that the motions of the heavenly bodies affect (or at least predict) mundane phenomena such as seasons and weather, floods and tides, and the generation of plants and animals. Secondly, and more significantly for our topic, uniform circular motion is a *mathematical* behaviour, which leads immediately to the modelling of the *Almagest*, while power to change the elements is by its nature not mathematical, since it operates with qualities such as hot and cold, wet and dry. And this creates a problem: how can cause-and-effect relations operating at the qualitative level, and largely within the "irregular" sublunary part of the cosmos, be well described by the predictive mathematical models of astrology?

[8] Robbins 1940.

It must be confessed that Ptolemy evades this problem. Essentially Ptolemy relies on the orderliness of the heavens to justify the mathematical structure of the predictive schemes of his astrology, but appeals to the disorderliness and complexity of the mundane environment to explain why astrological predictions, even when made according to the most correct principles, are not certain to be borne out. Moreover, the schemes that Ptolemy sets out to rationalise are in great part the rather chaotic traditional practices of the astrology of his time. Though he allows himself to reform or suppress some of this tradition in accordance with his physics, he can only go so far in that direction since his claim that astrology is a valid science depends heavily on the assumption that the traditional practices really work. One can sense his delight in finding here and there some apparent pattern in the jumble, for example when he finds harmonically significant ratios embedded in the "aspects" (astrologically significant linkages of zodiacal signs forming sides of triangles, squares, or hexagons), or when–indulging in a topos beloved of authors–he recovers from an old, neglected, and nearly illegible manuscript a gloriously complicated rationale for the seemingly nonsensical but empirically verified "Egyptian" system of terms (divisions of zodiacal signs associated with individual planets). Elsewhere Ptolemy almost seems to give up trying to explain, and lapses into catalogues of astrological associations scarcely distinguishable from the manuals of astrologers who were less sophisticated from a scientific point of view.

Optics, which in antiquity meant the study of visual perception, was a more fruitful subject for the interplay between mathematical and physical modelling. As in astronomy, there existed a range of well-established phenomena that lent themselves to explanation in terms of a geometrical model, in this case the "visual ray," diagrammed as a straight line extending from the viewer's eye to a point on an object. The hypothesis is that when a visual ray exists between the eye and a point on a body, that point is seen. The eye (or mind) always perceives the seen point as being in the direction in which the ray sets out from the eye, even if the ray is reflected or refracted at the interface between two bodies. This directional information provides the eye with indications of the shape, position, and movement of bodies; on the other hand, the ray conveys to the eye either limited knowledge or no knowledge at all about the distances to the point it perceives.

But what are these lines really? The classic exposition of Greek geometrical optics, repeatedly cited or paraphrased by later authors, was Euclid's *Optics*. This treatise does not explain the physical nature of the visual rays but does specify that they are discrete, with spaces between the individual rays that grow wider as the rays fan out towards more distant objects; moreover, some of the explanations of visual phenomena appear to assume that the rays are somehow attached to the eye (so that as the eye moves, the rays move accordingly). The gaps between the rays provide an explanation of the fact that objects are seen less clearly, or not seen at all, as they become more distant. But the gaps also lent themselves to a physical interpretation of the rays that is found in Peripatetic texts approximately contemporary with Euclid. According to this interpretation the eye emits, through pores in its surface, exquisitely thin and straight rods of matter (typically composed of fire) that extend with unimaginable swiftness until they encounter a body, at which point their progress may be impeded (in which case a visual perception of the body occurs) or reflected (if the surface is smooth enough) or refracted (if the body is porous).[9]

In his *Geography* Ptolemy treats the problem of planar map projections as essentially one of optics: how can one devise an appropriate framework of lines representing parallels and meridians to give the illusion of a part of a spherical surface?[10] Ptolemy takes for granted many elementary perspective consequences of the hypothesis of rectilinear visual rays. In one passage concerning the relation between the appropriate size of a map and the expected distance of the spectator from it, Ptolemy invokes the Euclidean gaps, which suggests that at this stage in his career (not long after the *Almagest* and *Tetrabiblos*), if he had any opinion at all about the physical nature of vision, it was not far removed from the Peripatetic notion of discrete material emanations.

When he came to write his *Optics* (a work that I suspect was among his last writings), Ptolemy had changed his mind.[11] He now speaks of the eye as emitting an entity conventionally translated as the "visual flux," a cone comprising a geometrical continuum of rectilinear rays that are stronger or weaker in perceptive power both to the extent that they have to extend a shorter or longer distance from eye to object, and to the extent that they are nearer to or further from the central axis of the cone.[12] It is to this weakening of the rays, rather than any supposed gaps between them, that fuzzy vision of distant or peripheral objects is due. Thus Ptolemy's geometrical treatment of visual phenomena thus preserves the parts of the Euclidean scheme that depend on the rectilinearity of the visual rays (namely, perspective phenomena, reflections, and refractions) but replaces the somewhat clumsy Euclidean handling of visual resolution with a more flexible and powerful hypothesis.

Unfortunately the entire first book of the *Optics* is lost, and with it Ptolemy's discussion of the physical makeup of the visual flux. Obviously he cannot have thought of it as a body, at least not the kind that displaces other bodies that formerly occupied its space, which is the only kind of body

[9]The Peripatetic texts and their possible relation to the Euclidean model are discussed in Jones 1994.

[10]Berggren and Jones 2000.

[11]The *Optics* survives, lacking its beginning and end, only in a medieval Latin translation of an Arabic translation, a circumstance that causes great difficulties of interpretation. The French translation in Lejeune 1989 and the English one in Smith 1996 are both useful, though under the circumstances neither can claim to represent Ptolemy's meaning exactly throughout.

[12]The Latin term rendered as "visual flux" is *uisus*, which almost certainly represents the same Greek word *opsis* that, in Euclidean optics, refers to the single visual rays; but Ptolemy used a different word when he meant an individual line of sight.

that is envisioned in Aristotle's or in Peripatetic physics. It seems likely that Ptolemy resorted to ideas from Stoic physics, which allowed for having distinct elements occupy the same space as if in layers. In this manner the Stoics could hypothesise that the entire cosmos was pervaded and regulated by *pneuma*. Ptolemy may have suggested that the eye issues the visual flux as an overlapping layer of matter in the space between eye and object; or perhaps more likely, he could have attributed to the eye a faculty of radiating a tensing power, creating the flux by means of the *pneuma* already present in the intervening space. We recall that in the *Planetary Hypotheses* he asserted a kinship between the motive power of the heavenly bodies and the analogous power in living things. As it happens, one of the very few references to Ptolemy's *Optics* in other authors that appear to pertain to its lost first book is a sentence in a work on physical topics by the eleventh-century Byzantine writer Simeon Seth: "Ptolemy says in his *Optics* that the visual *pneuma* is etherial and composed of the fifth element." This "visual *pneuma*" is probably the substance of the cone of the visual flux, and so we have a remarkable fusion of Aristotelian and Stoic element theory. The sixth-century philosopher Simplicius gives us a further clue when he writes:

> It should be noted that Ptolemy in his book *On the Elements* and in his *Optics*, and the great Plotinus, and Xenarchus in his *Difficulties Addressing the Fifth Element*, assert that motion in a straight line belongs to the elements when they are still in a place that is not natural to them, but (such motion) no longer belongs to them when they have assumed their natural place.... Manifestly they do not move when they are completely in their natural state, but, as the aforesaid men, i.e., Ptolemy, Xenarchus, and Plotinus, say, when they are in their natural state and in their proper places the elements either stand still or move in a circle.

This is precisely the notion that we have seen Ptolemy putting forward in the *Planetary Hypotheses*—which Simplicius does not cite here. It is obvious why Ptolemy would have repeated it in a (no longer extant) work on the elements; but what relevance can it have had in the *Optics*? I suspect that Ptolemy was invoking it here for a purpose converse to his purpose in the *Planetary Hypotheses*. There, the point was that etherial bodies in the heavens can spin freely and effortlessly even if their revolution is not concentric with the centre of the cosmos; in the *Optics*, perhaps Ptolemy claimed that the etherial matter of the visual flux, connected with our sight and thus *displaced* from its natural place in the heavens, travels in straight lines. A final observation worth making is that the Xenarchus cited by Simplicius as sharing this idea was active about the late first century B.C., so that here we may be able to identify the source of a principle that, in Ptolemy's hands, simultaneously accounts for the geometrical properties of the models of astronomy and of optics.[13]

I have saved for last what may have been Ptolemy's first major effort at mathematical modelling, the *Harmonics*.[14] The subject of this work calls for some explanation. Ancient Greek music was essentially melodic unison melody, occasionally employing singing or playing at the octave or the sounding of simultaneous distinct notes as an effect, but free of harmony in the modern sense. There existed numerous systems of relative pitches (i.e., scales) in which melodies could be composed, none of which involved a sequence of intervals quite like the diatonic scales on which most modern Western music is based. The science of harmonics, as Ptolemy presents it, investigates models that explain why certain intervals and combinations of intervals are esthetically pleasing and hence exist as constituents of the music actually produced in Ptolemy's time.

Unlike the kinematic models of the *Almagest* and the visual rays of the *Optics*, the models of the *Harmonics* are not geometrical but arithmetical. The model for any interval between musical pitches is a ratio of whole numbers, the question at issue being what rules determine the whole-number ratios that correspond to the intervals of existing musical scales. Ptolemy credits the ratio model to the Pythagoreans, though he disagrees with what he sees as their tendency to develop *a priori* modelling principles that are not referred to empirical evidence in an appropriate manner. In the course of criticizing the Pythagoreans (and the more fundamentally wrong-headed Aristoxeneans) and evolving his own models, Ptolemy makes more explicit pronouncements about the interplay between *a priori* and empirical reasoning in science than in any of his other works.

Ptolemy's harmonic models are built up from three kinds of esthetically satisfying intervals: (a) *homophones*, i.e., intervals between notes that sound nearly alike, being identical in pitch or separated by one or more octaves, modelled by ratios always of the type $m : 1$, e.g., $1 : 1$ or $2 : 1$ or $4 : 1$; (b) *concords*, i.e., intervals between notes that sound different but akin, and that form the more stable larger intervals in scales, modelled by ratios of the type $m : n$ such that m is often but not always equal to $n + 1$, e.g., $3 : 2$ or $4 : 3$ or $8 : 3$; and (c) the smaller *melodic* intervals between consecutive notes of a scale, which are almost always modelled by ratios of the type $(n+1) : n$, e.g., $9 : 8$.

The ratios are observable through the devices or instruments that make the notes. This is clearest in cases where the difference between notes follows from a difference between lengths in an instrument. For example, in wind

[13]The "fragments" of the lost part of Ptolemy's *Optics* (there are only four known) are collected in Lejeune 1989, 271. On Xenarchus, see Falcon 2001, 272, s.v. "Senarco."

[14]Barker 1989, 270-391 provides the best of the existing translations. West 1992 is a splendidly lucid introduction to all aspects of Greek music.

instruments one can measure the length of the pipe, say from the reed of an *aulos* (conventionally rendered by tin-eared classicists as “flute,” but actually a double reed like an oboe or shawm) to one of the finger-holes. For his harmonic demonstrations, Ptolemy prescribes instruments involving tensed strings, since these allow the maximum control and precision in the tunings and measurements. Thus it is by dividing a tensed string with a bridge into two parts in the ratio 4 : 3 that Ptolemy establishes the association of this ratio with the *tetrachord*, the principal fixed interval in the Greek scales (in modern terminology, a “fourth”).

But Ptolemy knows that length is not the only factor contributing to pitch. Thickness and density, among other characteristics of the bodies that produce the notes, are other variables that determine pitch; for this reason, before allowing us to try out ratios on a tensed string, Ptolemy instructs us to conduct a careful check of each part of the string to ensure that equal short lengths sound equal notes. Hence it is not at all easy to give a physical interpretation to the numbers in the modelling ratios that fully *explains* the musical intervals. Somehow a multiplicity of quantitative properties of a sounding body, some of them more straightforwardly measurable than others, give rise to a single abstract magnitude in the air in which the sound subsists.

In the chapters where he discusses the nature of sound and musical tone (*Harmonics* 1.3-4), Ptolemy does not try to explain the nature of sound more deeply than his initial definition that it is “a modification (*pathos*) of air when it is struck” (*Harmonics* 1.1), except for the conclusion that differences in pitch (“sharpness” and “heaviness”) are a form of quantity. He does, however, restrict the scope of harmonic science to the study of sequences of discrete sounds, each of which has a constant pitch, so that one may speak of stable relations or “ratios” between the notes. The special status of *whole-number* ratios enters the discussion circuitously, by way of the review of the Pythagorean model, and although Ptolemy uses divisions of a tensed string to provide empirical justification that the homophones and concords are modelled by ratios of small whole numbers, he provides no *a priori* justification of this fact.

But patience is rewarded. When Ptolemy has completed his set task of deducing a more or less complete set of models to describe the systems of tuning current in his time (*Harmonics* 3.2), he embarks on a new project of describing how harmonic theory illuminates our understanding of aspects of the cosmos that have no direct connection with sound, namely the behaviour of human beings and of the heavens. It turns out that harmonics is not really a science concerning sound at all. It is a science that discovers far deeper and more general truths about our world, exploiting one specific part of it that happens to be exceptionally well adapted to the interplay between sensory observation and rational deduction that, for Ptolemy, constitutes scientific method. The true subject of harmonics is *harmonia*, “the form of rational causation (i.e., causation arising from reason and intellect) that concerns good ratios of motions,” and this is necessarily present in all things that can move themselves, and above all in the most rational self-movers, namely, people and celestial spheres.

What this means is that the special status of whole-number ratios is a manifestation of the Good (in the Platonic sense) that the intellect apprehends and puts into action. One way that our intellects do this is by constructing musical instruments to produce sounds that fit the ideal ratios (since, after all, the sounds spontaneously produced by natural objects would not be recognized as music). Because of the close correspondence between measurable quantities in the instruments and the notes that we hear (which we can compare but not measure), we can discover the laws governing the order that our souls impose on this external matter. But these same laws are also recognizable, Ptolemy maintains, in the arrangement, motions, and powers of the heavenly bodies, which we discover through astronomy and astrology, and they must exist in our own characters, virtues, and emotions, where the quantitative relations are not apparent to our senses.[15]

These closing chapters of the *Harmonics* have received faint praise from modern readers, and it is undoubtedly true that the identification of detailed correspondences between the elements of his theory of musical tunings and an assortment of ethical, astronomical and astrological concepts is not Ptolemy's *forte*. But there can be no doubt that the principle motivating this *péché de jeunesse* was close to Ptolemy's heart, the conviction that the mathematical behaviour that we find here and there in the cosmos is structure imposed for the sake of the Good by minds upon a world that would otherwise be governed by disorder.

Bibliography

Barker, A. 1989. *Greek Musical Writings.* Vol. 2. *Harmonic and Acoustic Theory.* Cambridge.

Berggren, J. L., and A. Jones. 2000. *Ptolemy's* Geography: *An Annotated Translation of the Theoretical Chapters.* Princeton.

Dupuis, J. 1892. *Théon de Smyrne philosophe platonicien: Exposition des connaissances mathématiques utiles pour la lecture de Platon.* Paris.

Falcon, A. 2001. *Corpi e movimenti. Il* De caelo *di Aristotele e la sua fortuna nel mondo antico.* Napoli.

Goldstein, B. R. 1967. *The Arabic Version of Ptolemy's* Planetary Hypotheses. Transactions of the American Philosophical Society 57.4. Philadelphia.

Heiberg, J. L. 1907. *Claudii Ptolemaei Opera quae exstant omnia.* Vol. 2. *Opera astronomica minora.* Leipzig.

[15] See Swerdlow 2004 for a discussion of this part of the *Harmonics*, with emphasis on Ptolemy's celestial harmonics.

Jones, A. 1994. “Peripatetic and Euclidean Theories of the Visual Ray.” *Physis* 31, 47-76.

Lejeune, A. 1989. *L'*Optique *de Claude Ptolémée dans la version latine d'après l'arabe de l'émir Eugène de Sicile.* 2nd edition. Collection de travaux de l'Académie internationale d'histoire des sciences 31. Leiden.

Lloyd, G. E. R. 1978. “Saving the Appearances.” *Classical Quarterly* 28, 202–222. Reprinted with preface in G. E. R. Lloyd, *Methods and Problems in Greek Science*, Cambridge, 1991, 248-277.

Morelon, R. “La version arabe du *Livre des hypothèses* de Ptolémée.” *Mélanges de l'Institut dominicain d'études orientales du Caire* 21 (1993) 7–85.

Murschel, A. 1995. “The Structure and Function of Ptolemy's *Physical Hypotheses of Planetary Motion.*” *Journal for the History of Astronomy* 26, 33–61.

Neugebauer, O. 1975. *A History of Ancient Mathematical Astronomy.* 3 vols. Berlin.

Robbins, F. E. 1940. *Ptolemy: Tetrabiblos.* Loeb Classical Library 435. Cambridge, Mass.

Smith, A. M. 1996. *Ptolemy's Theory of Visual Perception: An English Translation of the Optics with Introduction and Commentary.* Transactions of the American Philosophical Society 86.2. Philadelphia.

Swerdow, N. M. 2004. “Ptolemy's *Harmonics* and the ‘Tones of the Universe’ in the *Canobic Inscription.*” *Studies in the History of the Exact Sciences in Honour of David Pingree*, ed. C. Burnett, J. P. Hogendijk, K. Plofker, and M. Yano. Islamic Philosophy Theology and Science: Texts and Studies 54. Leiden. 137–180.

Toomer, G. J. 1984. *Ptolemy's Almagest.* London.

__________ . 1987. “Ptolemy.” *Dictionary of Scientific Biography* 15, 207–224.

West, M. L. 1992. *Ancient Greek Music.* Oxford.