

# *Ptolemy's Search for a Law of Refraction: A Case-Study in the Classical Methodology of "Saving the Appearances" and its Limitations*

A. MARK SMITH

*Communicated by M. CLAGETT*

It is generally accepted that, after an auspicious beginning with EUCLID'S formulation of the ray-concept and an equally auspicious sequel, during which both *optics* proper and *catoptrics* were put on a firm scientific footing, the early development of mathematical optics was cut short by PTOLEMY'S failure to find the sine-law of refraction.<sup>1</sup> This failure, which apparently stemmed the growth of mathematical optics for some fifteen centuries, has long puzzled historians of science. We know for a start that, in attempting to derive the law experimentally, PTOLEMY was on the right track, and we also know that the experiment he devised

---

<sup>1</sup> The outlines of this "textbook" interpretation can be discerned in, or reconstructed from, a variety of standard sources: *cf., e.g.*, WILLIAM WHEWELL, *History of the Inductive Sciences*, second edition, vol. 2 (London, 1847), pp. 371–79; ERNST MACH, *The Principles of Physical Optics*, tr. J. S. ANDERSON & A. F. A. YOUNG (New York: Dover, 1953), pp. 39–40; RENÉ TATON, *Histoire générale des sciences*, vol. 2 (Paris: Presses Universitaires de France, 1957), pp. 299–301; VASCO RONCHI, *The Nature of Light*, tr. V. BAROCAS (London: Heinemann, 1970), pp. 108–112; and ALBERT LEJEUNE, *Euclide et Ptolémée: Deux stades de l'optique géométrique grecque* (Louvain: Université de Louvain, 1948), pp. 172–77. As far as actual influence is concerned, PTOLEMY'S *Optics* enjoyed remarkably little success until the Islamic Middle Ages, when it was used to some extent in AL-KINDI'S *De aspectibus* (ninth century) and to a very significant extent in IBN AL-HAYTHAM'S *Kitāb al-manāẓir* (c. 1000). The only version of the *Optics* now extant exists in a mid-twelfth century translation—which unfortunately lacks the first book and ends abruptly in the middle of the fifth book—by the Admiral EUGENE of Sicily from an undiscovered Arabic exemplar. Through this, and perhaps through other now-lost versions, the *Optics* became relatively well known in the Latin Middle Ages and remained so throughout the Renaissance. The first modern edition, drawn from a single manuscript, was published by GILBERTO GOVI in *L'Optica di Claudio Tolomeo* (Torino, 1885). Some seventy years later, ALBERT LEJEUNE published the currently accepted critical edition, *L'Optique de Claude Ptolémée* (Louvain: Université de Louvain, 1956); henceforth all references to the *Optics* will be from this edition. See pp. 27\*–37\* of LEJEUNE'S introduction for a good historical summary of the career of PTOLEMY'S *Optics*.

could have yielded the right results.<sup>2</sup> In short, the means of discovery available to PTOLEMY were essentially the same as those available to his seventeenth-century successors; yet where they triumphed, he failed. Why?

Two basic explanations have been offered. The simpler and more simplistic of these is that PTOLEMY's experimental technique was careless, from which it follows that, had he been more punctilious, he would have found the proper sine

<sup>2</sup> Echoing WHEWELL and GOVI, GEORGE SARTON extolls this as "the most remarkable experimental research of antiquity" (*Introduction to the History of Science*, vol. 1 [Baltimore: Williams and Wilkins, 1927], p. 274). Described by PTOLEMY in *Optics* V, 7-22, pp. 227-37, the actual experiment requires a hollow glass semicylinder (called a *baptistir*) that is closed at both ends. Into it a circular bronze template of the same diameter (figure 1) is inserted so that, when the *baptistir* is filled with water, the water's surface

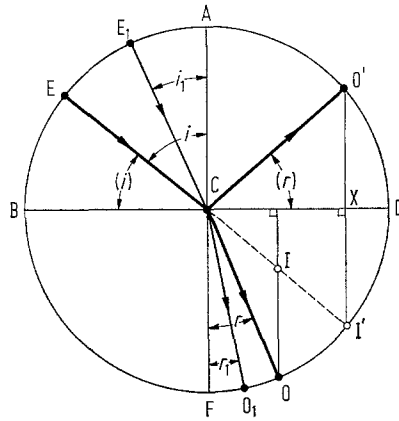


Fig. 1

will coincide with BD. The circumference of the template is then divided into gradations of  $1^\circ$ , a small marker is attached at center C, and a pointer E is set onto the circumference so that it can be moved along arc AB. With the template properly inserted into the filled *baptistir*, we can then place E at any position on arc AB, sight along line EC, and move a second pointer O along arc DF (below the water-level) until it falls into line with EC. That way we can find the corresponding  $r$  (measured by arc OF) for any  $i$  (measured by arc AE) from  $0^\circ$  to  $90^\circ$ . Using this same technique, apparently, PTOLEMY was able to measure  $i$  and  $r$  not only for air to water, but also for air to glass and water to glass. According to modern theory, of course, because the direction of radiation is opposite to that assumed in the visual ray theory, PTOLEMY was actually measuring refraction from the denser to the rarer rather than from the rarer to the denser medium. See PIERRE BRUNET & ALDO MIELI, *Histoire des science-santiquité* (Paris: Payot, 1935), pp. 824-33, for a relatively detailed treatment of PTOLEMY's refraction analysis that includes a French translation, based on GOVI's edition, of the relevant text of *Optics* V. For an English translation and analysis, also based on GOVI's edition, see MORRIS COHEN & I. E. DRABKIN, *A Source Book in Greek Science* (Cambridge: Harvard University, 1965), pp. 271-81. Far from perfect, this translation is nonetheless adequate to our basic needs and will therefore be used for direct quotations from PTOLEMY.

relation between  $i$  and  $r$ .<sup>3</sup> But this explanation rests on two very dubious assumptions: first, that such a relation is somehow ostensibly "in" the brute observations arising from his experiment, and second, that, even if it were, PTOLEMY could legitimately have been expected to see it. The alternative explanation is that, although he may not have uncovered the correct law, PTOLEMY did in fact stumble onto  $a$  law, which is implicit in his refraction tables and takes expression in the constancy of "second differences" to be found there.<sup>4</sup> While this certainly accounts for PTOLEMY'S failure to find the sine-law, it does not explain why he never acknowledged the law of second differences, or for that matter why he seems to have proposed no law whatever.<sup>5</sup> According to the one interpretation, then, being the consummate experimentalist he was, PTOLEMY could have—and indeed should have—discovered the sine-law but for some reason did not. According to the other, he had a law of sorts within his grasp but unaccountably failed or refused to accept it as such.

Against these two positions, I shall argue in this paper that PTOLEMY could not conceivably have found the sine-law and, furthermore, that he had sound reasons for not formulating a law based on constant second differences. The argument itself will hinge on the claim that, despite what both interpretations might lead us to believe, PTOLEMY did not undertake his refraction experiments with an eye towards discovery, because he already had a definite idea of what those experiments would tell him. He was not, in short, merely collecting and assessing bare facts with an open mind. On the contrary, long before he made his actual observations, PTOLEMY had every reason to expect that  $i$  and  $r$  would be constantly proportional (*i.e.*, that  $i : r :: i_1 : r_1$ ); and even after it was empirically confuted, this expectation subtly informed his search for a law of refraction.

In order to substantiate this claim, I shall first show that such an expectation was a logical consequence of PTOLEMY'S mathematical approach to optics. This will entail a general consideration of the theoretical foundations upon which his mathematical approach was based, as well as a specific consideration of how that approach was applied to certain fundamental optical problems. Then, against this background, I shall reconstruct PTOLEMY'S analysis of refraction, indicating precisely how it was affected by his presupposing a constant proportionality between  $i$  and  $r$  and, in the process, demonstrating that his failure to

---

<sup>3</sup> See, *e.g.*, GOVI, *L'Optica*, p. xxii. Throughout this essay,  $i$  will specifically designate the angle of incidence, whereas  $r$  will designate either the angle of refraction or the angle of reflection, depending on context.

<sup>4</sup> See the table on p. 232, which includes, in addition to the raw tabulations for  $i$  and  $r$ , two columns labelled  $d_1$  and  $d_2$ . The first of these columns gives the successive differences between  $r$ -values, and the second gives the differences between those differences. The constancy of these second differences ( $.5^\circ$ ) is obvious. Though easily inferred from his tabulations, this breakdown according to differences was not actually given by PTOLEMY but was imposed by GOVI (*L'Optica*, pp. xxiv–xxvii) and has been followed by virtually every commentator since. In the light of OTTO NEUGEBAUER'S work in ancient astronomy, it is clear that their basis in constant second differences puts these tabulations squarely within the Babylonian tradition of astronomical tabulations (*cf.* TATON, *Histoire générale*, vol. 1, p. 342).

<sup>5</sup> See, *e.g.*, LEJEUNE, *L'Optique*, p. 245, n. 33.

find a proper law was determined by that same presupposition. Finally, I shall point out how this failure reflects the shortcomings not of PTOLEMY, but of the theoretical basis underlying his analysis of refraction. Specifically, I shall show that the ray-concept upon which Euclidean-Ptolemaic optics depended was not, as is normally supposed, mathematically equivalent to the ray-concept as it has been understood since the seventeenth century—and, by extension, that the mathematical foundations of Euclidean-Ptolemaic optics were radically different from those of its seventeenth-century counterpart.

### The Theoretical Foundations

I have contended in a previous article that, as far as Greek science is concerned, “saving the appearances” was a definite methodological construct that applied not only to mathematical astronomy, but to mathematical optics as well.<sup>6</sup> The gist of my thesis was that the appearances-saving endeavor was carried out on the basis of four principal assumptions, the first and most fundamental of which is that all change or flux, insofar as it manifests irregularity, is just an “appearance” or illusion. From this follows the second supposition: that beneath the appearances there lies a real, intelligible world that is utterly simple, changeless, and eternal. This intelligible world, according to the third supposition, is a true Euclidean locus or “space”, within which things are *really* what they are by virtue of their spatial attributes and relationships. Within such a locus, moreover, the only real relationships are those most basic ones obtaining between and among points, and they are mathematically expressible in terms solely of rectilinear distances and angles. Fourth, these relationships are assumed to be immanent in, and thus immediately inferable from, the appearances.<sup>7</sup> In short, if properly understood, appearances do not really deceive us at all. They betoken a deeper and simpler reality that is rationally accessible to us in terms, finally, of straight lines and rectilinear angles.

Underlying this entire suppositional structure is the notion of a fundamental dichotomy between complexity and simplicity. Appearances, as manifestations of change and irregularity, are complex, while reality, as the manifestation of true changelessness, order, and uniformity, is simple. To “save” appearances is therefore to reduce them to the utter simplicity of uniformity. Such a reduction requires some absolutely simple and perfect gauge of uniformity, a *salvans*, according to which the appearances, or *salvanda*, fall ultimately into rational order. And what finally determines the perfection of the *salvans* is its conformance to what I have called the Principle of Natural Economy. This principle is the metaphysical pivot of the whole appearances-saving enterprise. Thus, perfect uniform circularity, which is the *salvans* of Classical mathematical astronomy,

---

<sup>6</sup> *Archive for History of Exact Sciences*, vol. 24, no. 2 (1981), pp. 73–99.

<sup>7</sup> This is tantamount to saying that the mathematical reality behind the objective appearances lies at a single level of abstraction from them, in the space within which they are manifested.

finds its pivotal justification in the fact that, of all isoperimetric figures, the circle and sphere are the most economical.

Granted these guiding assumptions and principles, the actual process of saving the appearances consists essentially of breaking up a given complex of *salvanda* (e.g., planetary motion in general) into simpler components. Each component within that complex constitutes an "anomaly" and, as such, must be reducible to uniformity, either directly or indirectly, according to the accepted *salvans*. If it is indirectly resolvable (and therefore still complex), it must be further subdivided until, at last, each sub-anomaly will dwindle to perfect uniformity. Consequently, the whole complex of *salvanda* comprises a nesting set of anomalies, in which the simpler is systematically subsumed under the more complex. Once brought to uniformity, any anomaly within that complex becomes a special case of the saving principle, and when the complex itself is brought to full uniformity, anomaly by anomaly, the appearances are saved.<sup>8</sup> Furthermore, the mathematical expression of that salvation will always be in the form of Euclidean proportions between straight lines and rectilinear angles—proportions whose terms are strictly spatial and represent real, underlying point-to-point relationships.

For instance, to all appearances retrograde motion is a matter of the vagaries of one point (planet P in figure 2) moving irregularly through space with respect to a fixed terrestrial viewpoint E. The anomaly itself consists in the fact that planet P seems on occasion to move backward, as is represented by continuous line  $P_1P_2$  to the left of the figure. According to the Ptolemaic system, however, this anomaly is a complex of two simpler elements and really involves three points, one of which—E, the center of the deferent—is fixed, and two of which—P and center C of the epicycle—move in relation both to E and to each other.

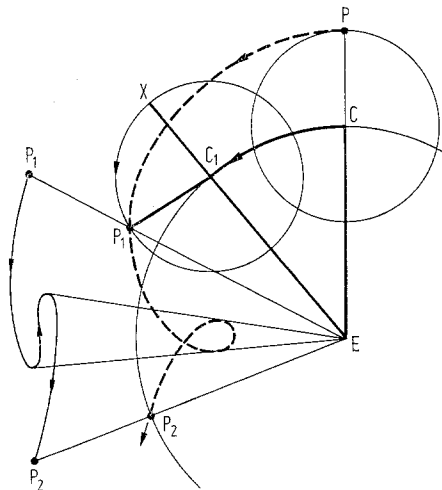


Fig. 2

<sup>8</sup> Thus, by systematically repacking these saved special cases, we should be able to reproduce the initial appearances exactly.

But P moves uniformly upon its epicycle, as indeed does C upon its deferent, so the anomaly is resolved into two perfectly uniform circular motions. In tandem, they will produce the complex, irregular orbit represented by the heavy broken line, and that orbit in its turn will produce the appearance of retrogression. Further, since the frequency and size of the retrograde loop are functions of the relative size of the epicycle, as well as the relative velocity of the planet upon it, the final mathematical salvation will reduce to the ratio of EC (the deferent's radius) to CP (the epicycle's radius) and arc  $CC_1$  (measured by  $\angle CEC_1$ ) to arc  $XP_1$  (measured by  $\angle XC_1P_1$ ). The final terms of salvation—straight lines and rectilinear angles—are therefore strictly spatial and represent point-to-point relationships that are really in, and immediately inferable from, the appearances.

### Saving *Optics*

As clearly as this analysis applies to Ptolemaic astronomy, it applies equally well, though perhaps less obviously, to Euclidean-Ptolemaic optics. To begin with, Euclidean-Ptolemaic optics, like its astronomical counterpart, has a single *salvans*—the visual ray. In essence, visual rays are just reified lines of sight projected in bundles from the eye to form a cone, whose apex defines the viewpoint and whose base marks out the visual field.<sup>9</sup> Each ray flows outward into that field until it strikes an optically opaque surface, which is thereby seen.<sup>10</sup> By establishing physical contact with object-surfaces, then, the visual ray mediates directly between us and the visible realm. More important, however, because that mediation is perfectly rectilinear, the visual ray establishes a direct spatial relationship between us and objects within the region carved out by the visual cone. This relation in turn depends on the shape, position, disposition, and size of those objects—in short, the compendium of mathematical attributes that makes them intelligible.<sup>11</sup>

Through its perfect, uniform rectilinearity, the visual ray provides us with a simple, rational way of perceiving these all-important attributes. Shape, for example, is perceived through points of intersection picked out upon the object-surfaces by incident rays, whereas position and disposition are perceived by means of ray-lengths and angles of incidence. And size-perception, finally, depends upon radially subtended angles at the viewpoint. Every visible surface within the visual field is thus transformed into a mosaic of intersection-points, each bearing a unique, fixed linear and angular relation to the viewpoint. The sum total of such

<sup>9</sup> EUCLID, *Optics*, defs. 1 and 2, in *Euclidis opera omnia*, vol. 7, ed. I. L. HEIBERG (Leipzig: Teubner, 1895), p. 2. All subsequent references to EUCLID's *Optics* will be from this edition.

<sup>10</sup> EUCLID, *Optics*, def. 3, p. 3.

<sup>11</sup> The visual ray theory therefore rests on an implicit distinction between the mere *visibilia*—the properties (or property) that make external objects accessible to visual sensation—and the *perceptibilia*—those properties that make them accessible to mental scrutiny. ARISTOTLE seems to acknowledge this distinction explicitly with his “common sensibles”—*i.e.*, movement, rest, figure, magnitude, number, and unity—all of which are mathematical or mathematically based (*De anima* III, 1, 425a14–17).

relations, mediated throughout by the visual ray, tells us everything worth knowing about visible objects: namely, what and where they really are in respect to us within the Euclidean locus that constitutes the true, intelligible world.

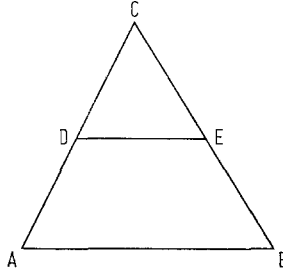


Fig. 3

To make this point more obvious, let us use theorem 21 of EUCLID'S *Optics* as an example.<sup>12</sup> In that theorem EUCLID poses the following problem: assuming that we see a magnitude of unknown length (AB in figure 3) from an indeterminate location C, to find its real length. The heart of the problem is that AB will *appear* to be smaller than it really is, its relative smallness depending on the distance of C from AB and therefore on the size of subtending angle ACB. To "save" this appearance, we draw rays CA and CB connecting the viewpoint with the endpoints of the magnitude, place known length DE parallel to AB at such a position that it appears to be the same length (*i.e.*, so that its endpoints are touched by the rays), and measure CD and CA. By Euclidean geometry we know that, since DE and AB are parallel,  $\angle CDE = \angle CAB$ , and  $\angle CED = \angle CBA$  (which is to say that  $\angle CDE : \angle CED :: \angle CAB : \angle CBA$ ). Therefore, we know that  $CD : DE :: CA : AB$ . Having measured CD and CA, and knowing the length of DE, we can thus determine the true length of AB. That is, we have saved the appearance of smallness by determining true size; and the same approach can be used to save a wide variety of optical appearances, including, for example, the fact that a large circle seen at a certain distance and obliquity will seem to be what we would now designate as a small near-ellipse.<sup>13</sup>

Three significant features link this optical example to the previous astronomical one. First, the salvation ultimately comes to ground in Euclidean proportions between rectilinear lengths (*i.e.*,  $CD : ED :: CA : AB$ ) and angles ( $\angle CDE : \angle CED :: \angle CAB : \angle CBA$ ). Second, the linear and angular terms of these proportions represent direct spatial relations between viewer and viewed, relations that are immediately inferred from the phenomena. Third, and perhaps most significant, the salvation itself hinges on an absolutely simple *salvans*, the visual ray, whose saving grace is its perfect, uniform rectilinearity. So, as far as mathematical optics is concerned, the visual ray is *essentially* a straight line, and a straight line is the measure of perfect point-to-point brevity. In short, the *salvans* for optics, like that for astronomy, conforms to the Principle of Natural Economy.

<sup>12</sup> *Optics*, p. 12.

<sup>13</sup> Cf. EUCLID, *Optics*, props. 34–36, pp. 60–80.

There is yet one further link between Euclidean-Ptolemaic optics and Ptolemaic astronomy, the crucial nature of which will become evident in the following section. By PTOLEMY'S time, the domain of *salvanda* in optics, like that in astronomy, had been resolved into a definite sequence of anomalies, determined in ascending order of complexity by whether the visual ray is unbroken, fully broken or reflected, or only partially broken or refracted—hence, the three traditional subdivisions of *optics*, *catoptrics*, and *dioptrics*.<sup>14</sup> According to this structure, then, Euclidean-Ptolemaic optics evolves in a logical succession of salvations, starting with the very simple “anomaly” of direct vision, progressing to the more complicated anomaly of reflection, and culminating with the even more complicated anomaly of refraction. Within this succession, the difference between anomalies/salvations is one of degree rather than kind, so that reflection is presumably just a complex form of direct vision, as is refraction of reflection. To see how this presumption was vindicated in practice, we have but to look at the classical analysis of reflection.

### Saving *Catoptrics*

While the primary concern of *optics* is the disparity between what a given object looks like and what it actually is in space, the overriding concern of *catoptrics* is the disparity between where the object seems to be and where it actually is in space. In other words, the fundamental problem or appearance of reflection is that of image-location, and the key to saving it lies in the relationship of equality between the angles of incidence and reflection.<sup>15</sup> Recognized quite early—probably

---

<sup>14</sup> This analytic structure is quite clearly reflected in the organization of PTOLEMY'S *Optics*, which is divided into three major, and absolutely distinct, sections. The first of these, which ends with book II, treats direct vision, starting with first principles and culminating with the analysis of binocular vision. The second section, which treats reflection and comprises books III and IV, starts with the analysis of plane mirrors, passes to the more complex case of convex mirrors, progresses to the even more complex case of concave mirrors, and ends with compound mirrors. Finally, in the third section, which consists of book V, PTOLEMY undertakes his analysis of refraction. Even the most perfunctory reading of the *Optics* will show that, within the general confines of this basic structure, PTOLEMY was trying, insofar as possible, to build the science of optics, step by theorematic step, as EUCLID had built the science of geometry.

<sup>15</sup> In figure 1, n. 2 above, if BD is taken to represent a plane mirror, EC the incident ray, C the point of reflection, O' the visible object, and CO' the reflected ray, then (i) will always be equal to (r). Moreover, since the image I' of object O' lies at the juncture of  $\perp O'X$  (the cathetus) and extension CI' of incident ray EC, and since (i) =  $\angle DCI'$  = (r), then I' must lie precisely the same distance along cathetus O'X below BD as O' lies along the cathetus above BD—i.e.,  $O'X = XI'$ . The same sort of analysis also enables us, *mutatis mutandis*, to locate I' in convex and concave mirrors. On the basis of these principles, PTOLEMY was able to derive the Law of Equal Angles experimentally in *Optics* III, 7–12, pp. 91–94, using the same template designed for the refraction experiments. In this case, disk ABFD in figure 1 is laid horizontally, and to it are applied, in succession, a plane mirror, a convex mirror, and a concave mirror, such that all three reflecting surfaces will be vertical to ABFD, and BD will either coincide with, or be tan-



by EUCLID and certainly by ARCHIMEDES— this all-important relation was accepted for at least three centuries as an empirical fact and, as such, carried the force of probability if not necessity. That, however, changed dramatically nearly a century before PTOLEMY, when HERO of Alexandria transformed mere empirical likelihood into law with a simple but ingenious mathematical proof.

Assume, he says in proposition 4 of the *Catoptrics*,<sup>16</sup> that the resulting angles are not equal. For instance, let EH in figure 4 represent a plane mirror, G the eye, and D the object; and let ray GB strike the mirror at B and reflect to D such that  $\angle GBE$  of incidence  $\neq \angle DBH$  of reflection. Meantime, draw rays GA and AD so that  $\angle GAE = \angle DAH$ . Under these conditions, HERO concludes, it can easily be proved that the sum of GA and AD is invariably less than the sum of any other lines connecting G and D to any point on the mirror's surface other than A.<sup>17</sup>

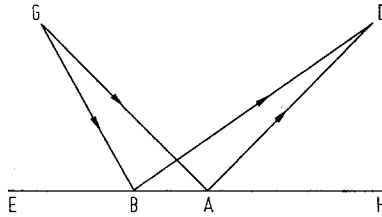


Fig. 4

With this seemingly trivial proof, what HERO did in effect was to demonstrate the *necessity* of the relation of equal angles by recourse to the so-called Principle of Least Lines. That is, he proved that the path forged by those rays subtending equal angles with the mirror's surface is the shortest possible and therefore the most economical. Consequently, he actually resolved the Law of Equal Angles, the pivotal saving principle of reflection, into a special case of perfect point-to-point brevity, the fundamental *salvans* of *optics*. In accomplishing that, furthermore, he demonstrated that the overall anomaly of reflection, with all its attendant sub-anomalies, is merely a special case of unimpeded visual radiation, itself a special case of the Principle of Natural Economy. *Optics* and *catoptrics* were thus irrevocably drawn by Hero into the same systematic pale. This in turn dissolved any real difference, other than in complexity, between direct vision and reflection, both anomaly-types having been mathematically reduced to absolute order through the same saving principle.

gent to, the surface of reflection at C. Then, sighting along EC, we move pointer O' until it appears to fall into line with EC, and in all three mirrors (*r*) will invariably prove to be equal to (*i*). It is hardly surprising that, coupled with his refraction experiments, these reflection experiments should be adduced as clear evidence of PTOLEMY'S strong empirical bias in optics.

<sup>16</sup> *Heronis Alexandrini opera quae supersunt*, vol. 2, fasc. 1, ed. WILHELM SCHMIDT & L. NIX (Leipzig: Teubner, 1900), pp. 324–29.

<sup>17</sup> HERO gives an equivalent proof for convex mirrors in *Catoptrics*, prop. 5, pp. 328–31. I should perhaps point out that it was common practice among the Greeks, as well as among medieval optical writers, to measure the angles of incidence and reflection with respect to the surface of reflection rather than the normal.

Perhaps even more important than explicitly validating the Law of Equal Angles, though, HERO'S proof also implicitly validated the entire analytic structure of Euclidean-Ptolemaic optics. For, by specifically showing that reflection amounts to nothing more than a complex form of direct vision, his proof raised the more general likelihood that, as an anomaly, every optical appearance, as well as its salvation, is just a more or less complex form of some simpler, more basic type. In short, he provided an explicit and exemplary justification of the heuristic according to which the science of Euclidean-Ptolemaic optics seems to have developed from its inception. Hence, in addition to vindicating the saving principle of reflection, HERO'S proof marked a critical step in the establishment of a firm deductive basis, according to a clearly defined and totally interdependent system of anomalies, upon which to ground the eventual salvation of all possible optical appearances.<sup>18</sup>

As far as our understanding of PTOLEMY'S methodology is concerned, this point has a twofold bearing. First, it puts his much-touted reflection experiment into clear perspective, not as a means of discovering fact or testing hypothesis, but simply as a way of graphically verifying what he already knew: a law whose force had been established nearly a century before on purely deductive grounds.<sup>19</sup> Thus, in this important case at least, experimentalism dwindles to relative insignificance as a factor in PTOLEMY'S optical analysis—which strongly suggests that he considered induction to be subordinate to deduction. Second, being heir to the same optical tradition as HERO, a tradition to which HERO himself had contributed significantly, PTOLEMY was heir to the same heuristic. Methodologically this is critical, because it means that, guided by that heuristic, he was logically impelled to suppose that, as part of a well-defined system of anomalies, refraction would prove to be a special case of reflection, just as reflection had proved to be of direct vision. Hence, in approaching refraction he was already predisposed to treat it as systematically related to, rather than fundamentally different from, reflection.

### Trying to Save *Dioptrics*

Just how closely related PTOLEMY actually thought reflection and refraction were is evident from his analysis of the similarities between them. "Visual rays," he informs us, "may be altered in two ways: (1) by reflection, i.e., the rebound from objects, called mirrors, which do not permit of penetration, and (2) by bending in the case of media which permit of penetration. . . ."<sup>20</sup> Both alterations, in other words, are functions of the ray's striking a surface and then being deflected along another path. The two are also alike, he continues, in that they entail

<sup>18</sup> Thus, all optical problems, from those involving unbroken visual radiation to those involving refraction, ought ultimately to be resolvable through the Principle of Least Lines.

<sup>19</sup> Obviously, then, PTOLEMY was not relying upon experiment to discover the relationship of equal angles, nor was he employing it as an *experimentum crucis* in the Baconian/Newtonian sense.

<sup>20</sup> *Optics* V, 1, p. 223, lines 2–7, tr. COHEN & DRABKIN, *Source Book*, pp. 271–72.

“the relation of images and objects,”<sup>21</sup> which is to say that in each case the basic anomaly consists in the disparity between real and apparent location. In both cases, therefore, saving the appearances involves a mathematical reconciliation of these loci. Finally, PTOLEMY claims that in both reflection and refraction, not only do incidence and deflection occur in the same plane, but the image is located in that plane at the juncture of the line of incidence and what amounts to the cathetus.<sup>22</sup>

Despite these similarities, however, reflection and refraction differ in two signal ways according to PTOLEMY. First, while reflection involves the ray's striking a body with an optically impenetrable surface, refraction involves its striking a body with an optically permeable surface. But the permeability of such a body is contingent on its relative “density,” so when a ray passes from a relatively permeable/rare to a relatively impermeable/dense medium, it will be bent toward the normal. Conversely, passage from a dense to a rare medium will cause an equivalent bending away from the normal.<sup>23</sup> In either case, the greater the density-differential, the greater the bending.<sup>24</sup> The second difference is that, unlike reflection, refraction “does not take place at equal angles,” although PTOLEMY is quick to add somewhat cryptically that “the angles, as measured from the perpendicular (*i.e.*, the normal) do have a definite quantitative relationship.”<sup>25</sup>

To construe this unspecified relationship as one of constant proportionality between *i* and *r* is indeed tempting, as COHEN & DRABKIN remark.<sup>26</sup> More than tempting, however, it is almost unavoidable in light both of PTOLEMY's explicit acknowledgment of the close ties between reflection and refraction and of his evident commitment to the idea that refraction marks the last and most complex in a sequence of three anomaly-types, the first two of which had long since been

<sup>21</sup> *Optics* V, 1, p. 223, lines 10–12, tr. COHEN & DRABKIN, *Source Book*, p. 272

<sup>22</sup> Cf. *Optics* V, 3–5, pp. 224–26. In the first of these claims PTOLEMY is simply stating that the center of sight, the point of reflection/refraction, the object-point, and the image all lie in the same plane. Thus, in figure 1, n. 2 above, E, C, O', and I', as well as E, C, O, and I (the image-point in refraction), all lie in plane ABFD. Using the same figure, we can see that the image in refraction, I, like that in reflection, I', is located at the juncture of EC extended and the perpendicular dropped to the surface of refraction.

<sup>23</sup> Although he hints at it in his preliminary discussion in V, 2, p. 224, PTOLEMY does not give a definitive statement of this principle of reciprocity until the end of his experimental analysis, in V, 31, p. 243.

<sup>24</sup> *Optics* V, 31, p. 243 and V, 33, p. 244. By relating refractivity to the density-differential, PTOLEMY was of course adumbrating the modern conception of an index of refraction. An interesting implication of his assuming that the greater the density-differential, the greater the bending, is that, in passing into increasingly dense media, the ray will be increasingly bent toward the normal until, finally, it will theoretically be forced to coincide with the normal. At this point, however, the density will have become so great as to render the medium optically impermeable, and reflection will ensue. So in a sense reflection is merely refraction carried to the extreme.

<sup>25</sup> *Optics* V, 2, p. 224, lines 6–8, tr. COHEN & DRABKIN, *Source Book*, p. 272. Actually, this statement is a good deal more cryptic in Latin than the English translation might lead us to believe (... nulla fit in eis flexio ad equales angulos, sed habent similitudinem quandam et quantitatem que sequitur habitudinem perpendicularium ...).

<sup>26</sup> *Source Book*, p. 272, n. 2.

reduced to equivalency through the least-lines principle. Thus, since the salvation of reflection had proved to be no more than a systematic variant of the salvation of direct vision, it stood to reason that the salvation of refraction would in turn prove to be a systematic variant of the salvation of reflection, especially since the two phenomena display such obvious affinities. In tackling refraction, then, it was only natural for PTOLEMY to take reflection as his analytic model. But if reflection was saved through a form of constant proportionality—equality—between  $i$  and  $r$ , then why not refraction too? Why not, in short, draw the obvious conclusion that the governing principle of refraction must be equivalent to, but more complex or general than, the Law of Equal Angles—*i.e.*, that  $i : r :: i_1 : r_1$ ?<sup>27</sup>

Assuming that he actually did follow this self-evident train of deductions, PTOLEMY must have undertaken his refraction experiments expecting at the outset only to verify the constant proportionality that reason led him to anticipate. So much, at least, is consistent with his use of experiment in reflection. In this case, though, in lieu of confirmation, experiment yielded confutation, because, of course, there is no constant proportionality between  $i$  and  $r$  in refraction. In fact, so disproportionate are they, even according to PTOLEMY's results, that he could not possibly have found one without denying the evidence of his senses. It would therefore seem that, in the face of clear facts, he was forced to repudiate his preconceived "law" of constant angular proportionality; yet, as will become clear in a moment, rather than abandon that law, he simply adjusted it to the observations and, in the process, adjusted the observations themselves in a very telling way.

Take his results for refraction from air to water, reproduced in the table below:<sup>28</sup>

$i$	$r$	$d_1$	$d_2$
0	0		
10	8	8	.5
20	15.5	7.5	.5
30	22.5	7	.5
40	29	6.5	.5
50	35	6	.5
60	40.5	5.5	.5
70	45.5	5	.5
80	50	4.5	.5

Clearly, this table is laid out so that  $i$  occurs in a uniform progression from  $0^\circ$  to  $80^\circ$  at a constant rate of  $10^\circ$ . Against this uniform progression, we have the apparently non-uniform progression of  $0^\circ, 8^\circ, 15.5^\circ, 22.5^\circ, \dots, 50^\circ$  for  $r$ . However, as is evident from the  $d_1$  column, the first differences between  $r_1, r_2, r_3,$

<sup>27</sup> This conclusion rests on the obvious fact that  $i = r$  is just a specific form of the more general proportionality  $i : r :: i_1 : r_1$ .

<sup>28</sup> PTOLEMY's tabulations, upon which this table is based, are found in *Optics* V, 11, pp. 229–30. I reiterate that the  $d_1$  and  $d_2$  columns were not supplied by PTOLEMY; they are modern interpolations.

and so forth, decrease constantly by a factor  $d_2 = .5^\circ$ , the second difference. In other words, for every constant increase in  $i$ , there is a corresponding decrease in  $d_1$ . If we take the first  $d_1$  (which is also the first value for  $r$ —i.e.,  $8^\circ$ ) as a base factor, however, and suppose that, instead of decreasing, it remains constant, we will get the uniform series  $8^\circ, 16^\circ, 24^\circ, 32^\circ, \dots 64^\circ$  for  $r$ , which, plotted against the  $10^\circ, 20^\circ, 30^\circ, 40^\circ, \dots 80^\circ$  series for  $i$ , yields the originally anticipated "ideal" proportionality  $i:r :: i_1:r_1$ . In reality though, after an initial sweep of  $8^\circ$ , the base factor decreases uniformly at a rate of  $.5^\circ$  for every  $8^\circ$ , so that the final result is  $0^\circ + 8^\circ = 8^\circ, 8^\circ + 7.5^\circ = 15.5^\circ, 15.5^\circ + 7^\circ = 22.5^\circ, 22.5^\circ + 6.5^\circ = 29^\circ$ , etc. Therefore, the actual  $r$ -series is a function of two basic factors: an "ideal" uniform progression according to constant increments of  $8^\circ$  and a uniform regression of those increments according to constant decrements of  $.5^\circ$ .<sup>29</sup>

<sup>29</sup> This is easily understood if we translate the problem into terms of motion, with the  $i$ -series representing a uniform circular motion at a constant rate of  $10^\circ/\text{time-unit}$ . Then, as  $i$  moves uniformly from  $0^\circ$  to  $80^\circ$ ,  $r$  would ideally move from  $0^\circ$  to  $64^\circ$  at a constant rate of  $8^\circ/\text{time-unit}$ ; and the two motions would be constantly proportional. According to observation, however, instead of moving uniformly,  $r$  seems to decelerate constantly, so that, instead of sweeping out the ideal arc of  $64^\circ$ , it sweeps out an actual arc of  $50^\circ$ . The problem thus reduces to this: how can we make the ideal uniform motion of  $r$  through an arc of  $64^\circ$  at a constant rate of  $8^\circ/\text{time-unit}$  appear to be a non-uniform motion through an arc of  $50^\circ$  at a rate that changes constantly by decrements of  $.5^\circ/\text{time-unit}$ ? A very close, and very simple, first approximation to "saving" this appearance can be found in the Ptolemaic eccentric-model. Let AB in figure 5 represent a circular orbit, whose center is C, and let the ideal uniform motion of  $r$  be plotted on the circle from point C such that  $r$  sweeps out arc DK of  $64^\circ$  in constant increments of  $8^\circ$  (i.e., arcs DF, FG, GH, etc. =  $8^\circ$ ). Then pick an eccentric point E from which arc DK will appear to be  $50^\circ$  and from which arcs DF, FG, GH, etc. will appear to be  $8^\circ, 7.5^\circ, 7^\circ$ , etc. respectively. Thus, as  $r$  sweeps out the uniform succession of  $8^\circ, 16^\circ, 24^\circ, 32^\circ$ , etc. measured from C, it will seem from point E to sweep out the non-uniform series  $8^\circ, 15.5^\circ, 22.5^\circ, 29^\circ$ , etc.—the values actually given in PTOLEMY'S table. It should be noted, though, that without a number of subtle refinements, this simple eccentric-model will not save the appearances exactly, because the rate of apparent deceleration in  $r$  is not perfectly constant at  $.5^\circ/\text{time-unit}$  but continually decreases as  $r$  approaches K.

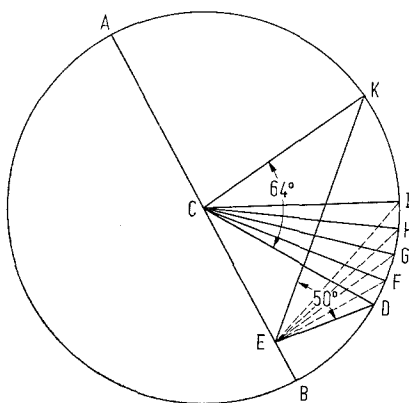


Fig. 5

This analysis extends to PTOLEMY's two remaining refraction tables as well. In the first of these (air to glass) the base factor is  $7^\circ$ , and in the second (water to glass) it is  $9.5^\circ$ , but in both—as in the table for air to water—the second difference remains unchanged at  $.5^\circ$ .<sup>30</sup> In all three cases PTOLEMY reduced the

<sup>30</sup> Given in *Optics* V, 18, pp. 233–34 and V, 21, pp. 236–37, PTOLEMY's results for refraction from air to glass and water to glass are as follows:

air-glass		water-glass	
<i>i</i>	<i>r</i>	<i>i</i>	<i>r</i>
0	0	0	0
10	7	10	9.5
20	13.5	20	18.5
30	19.5	30	27
40	25	40	35
50	30	50	42.5
60	34.5	60	49.5
70	38.5	70	56
80	42	80	62

That PTOLEMY's adjustments were indeed purposeful and not merely due to random error is evident from a comparison of his values with the correct values according to the sine-law. For instance, compare his values against the values of *r* for refraction from air to water, taking 1.33 as the index of refraction and rounding off to the nearest  $.5^\circ$ —the probable limit of his ability to discriminate:

<i>i</i>	<i>r</i> (Ptolemy)	<i>r</i> (modern)
0	0	0
10	8	7.5
20	15.5	15
30	22.5	22
40	29	29
50	35	35
60	40.5	40.5
70	45.5	45
80	50	47.5

Notice, first, that the adjustments are selective. Three of the eight Ptolemaic values are accurate, and the remaining five are all adjusted upward. Second, notice that, with one signal exception, these adjustments are all uniform at  $.5^\circ$ . Moreover, in that one exceptional case (when  $i = 80^\circ$ ), the discrepancy between *r* (PTOLEMY) and *r* (modern) is unusually large ( $2.5^\circ$ , well beyond the  $.5^\circ$  limit of error)—too large and too idiosyncratic to be attributed to mere observational error. Surely, then, the likelihood of PTOLEMY's having made these particular adjustments by pure chance to home in on this particular rational sequence of *r*-values is minimal. It is of course remarkable that, even after adjustment, PTOLEMY's values are for the most part exceptionally close to the ones he "should" have found.

anomaly—*i.e.*, the disparity between the observed and ideal or anticipated value of  $r$ —to two absolutely simple components: a uniform progression according to a constant base-rate ( $7^\circ$ ,  $8^\circ$ , or  $9.5^\circ$ ), and a uniform regression of that base-rate at a constant rate of  $.5^\circ$ . He therefore seems to have pinned his hopes for saving refraction on establishing a simple mediate relation between  $i$  and  $r$  through the constant proportionality of  $i$  and  $R$  (ideal or anticipated  $r$ ) and a constantly variant relation between  $R$  and  $r$ .<sup>31</sup>

### In Search of a Law

By skewing his data in the way he did, PTOLEMY made it quite easy to reach a general solution for  $r$  in terms of  $i$ ,  $R$ , and  $d_2$ . In the first place, it is evident from all three tables that, for any value of  $i$ ,  $r$  is equal to the sum of  $d_1$ 's up to that point, and that sum differs from the value of  $R$  by the overall amount of decrease in the  $d_1$ 's. In the air-water table, for example, the value of  $r$  for  $i = 20^\circ$  is  $8^\circ + 7.5^\circ = 15.5^\circ$ , but if  $d_1$  were to remain constant at the ideal rate of  $8^\circ$ , while  $i$  progresses from  $0^\circ$  to  $20^\circ$ ,  $r$  would be  $16^\circ$ . So there is a difference of  $.5^\circ$  ( $1 \times d_2$ ) between  $R$  and  $r$ . Similarly, the value of  $r$  for  $i = 30^\circ$  is  $8^\circ + 7.5^\circ + 7^\circ = 22.5^\circ$ , which differs by  $1.5^\circ$  ( $3 \times d_2$ ) from the  $R$  of  $24^\circ$ , whereas for  $i = 40^\circ$ ,  $r = 29^\circ$ , which differs by  $3^\circ$  ( $6 \times d_2$ ) from the  $R$  of  $32^\circ$ . From this we can conclude that, for any  $i$  whatever, the difference between  $r$  and  $R$  will be  $\sum_1^{n-1} d_2$ , where  $n = i/10$ .<sup>32</sup>

In short,  $r = R - \sum_1^{n-1} d_2 = R - \frac{n(n-1)}{2} d_2$ . Thus, when  $i = 70^\circ$ ,  $n = 7$ ,  $R = 7^\circ \times 8^\circ = 56^\circ$ , and  $r = 56^\circ - \frac{7 \times 6}{2} \times .5^\circ = 45.5^\circ$ , which is indeed the value given in the table. With this same technique we can compute  $r$  for any value of  $i$  we please, so that, for example, to find  $r$  for  $i = 1^\circ$ , we compute  $n$  (which equals  $1/10$  or  $.1$ ) and  $R$  (which is  $n \times$  base-rate  $8^\circ$  or  $.8^\circ$ ), and put those values into the equation to get  $r = .8^\circ - \frac{.1 \times (-.9)}{2} \times .5 = .8225^\circ$ .<sup>33</sup>

It should be clear that, besides providing a general solution for  $r$ , this equation also offers a clear, concise, and explicit statement in modern guise of the law of second differences implicit in PTOLEMY's analysis of refraction. First, it follows

<sup>31</sup> Ideally, then,  $R$  ought to differ from  $r$  by some constant factor  $x$  (presumably  $= .5^\circ$ ), which varies constantly with  $i$  or  $R$  so that, if  $R - r = x$ , then  $a(R - r) = ax$ . In such a case,  $i : (r + x) :: ai : a(r + x)$ , which is a relatively simple permutation of  $i : r :: i_1 : r_1$ .

<sup>32</sup> This value for  $n$  follows from our taking the value  $r = 8^\circ = R$ , when  $i = 10^\circ$ , as the standard or ideal from which all other values will derive or deviate. Thus, for any  $n$ -multiple of  $i$ ,  $r$  will be the  $n$ -multiple of  $8^\circ$  minus the summation-factor.

<sup>33</sup> This method can be applied to the other two refraction tables as well. Thus, to compute  $r$  for  $i = 1^\circ$  in the case of refraction from air to glass (where the base-value of  $R$  is  $7^\circ$ ), we get the equation  $r = \frac{7^\circ}{10} - \frac{.1 \times (-.9)}{2} \times .5^\circ = .7^\circ + .0225^\circ = .7225^\circ$ .

directly from his tabulations and is therefore immediately “in” the phenomena as he recorded them. Second, it conforms exactly to the proposed model of uniform progression countervailed by constant regression in the rate of progression. Third, it is absolutely general. Finally, and most telling, it is wholly contingent on  $d_2$ .<sup>34</sup> Thus, by approaching the problem as he apparently did, and skewing his results accordingly, PTOLEMY had within his reach a definite and almost intuitively obvious law of refraction based on second differences, the very thing for which he was supposedly searching. Nonetheless, as his silence regarding that law would appear to indicate, he somehow let it slip through.

At first glance, this apparent lapse might seem startling, but if we consider the problem as PTOLEMY posed it for himself—to establish a direct relation between  $i$  and  $r$  in the restricted form of Euclidean proportions—then it becomes readily understandable. For even if PTOLEMY had found the general solution at a practical, computational level, the difficulty of expressing it satisfactorily in the necessary  $i : r$  format would have been virtually insuperable. For instance, it was doubtless as clear to PTOLEMY as it is to us that the relation between  $r$  and  $R$  involves a continual variation in the difference between the two. That is, as  $r$  and  $R$  increase, so does the difference between them. Moreover, it is equally clear that this increase is a function of  $d_2$ . The problem is to determine exactly what that variation is—in other words, precisely what sort of function of  $d_2$  it is.<sup>35</sup>

To us, this poses no difficulty whatever. We define the variation, and thus give it definite expression, in the algebraic form  $\frac{n^2 - n}{2} d_2$ , which involves a simple quadratic function divided by 2 and multiplied by  $d_2$ . To PTOLEMY, on the other hand, such a formulation would have been incomprehensible for a variety of reasons. To start with, in Greek mathematics, fractions represent ratios, not numbers, so  $\frac{n^2 - n}{2}$  would translate into  $(n^2 - n) : 2$ . Not only is such a ratio peculiar; it is senseless in Euclidean terms, because it entails a direct relation between a linear and an areal magnitude. Worse, the quadratic function  $n^2 - n$  itself involves a direct relation, through subtraction, between a linear and an areal magnitude. Worse yet, this whole complex of meaningless relations is directly related, through multiplication, to  $d_2$ . Finally, add to all these complications the further complication of relating this extraordinarily complex set of relations directly to  $r$  and then relating *that* relationship directly to  $i$  in order to get the requisite

<sup>34</sup> GOVI, *L'Optica*, p. xxvi, offers an entirely different solution in the algebraic form  $r = ai - bi^2$ , where  $a$  and  $b$  are constants that depend on the table (e.g., in the case of air to water,  $a = .825$  and  $b = .0025$ , whereas in the case of air to glass,  $a = .725$  and  $b = .0025$ ). Although perfectly adequate, this solution is not intuitively obvious from the data, nor does it show any obvious formal contingency on  $d_2$ . In short, it is too abstract, too removed from the phenomena, to have been evident to PTOLEMY.

<sup>35</sup> The nub of the problem is that, while  $i$  increases arithmetically, the difference between  $R$  and  $r$  increases geometrically so that, rather than a constant variation, there is a constantly increasing variation. This of course dashes any hope of establishing a constant “mediate” proportionality of the kind suggested in n. 31 above.



constant proportionality  $i:r + \frac{n^2 - n}{2} d_2$ , and the magnitude of PTOLEMY'S problem becomes clear.<sup>36</sup>

Small wonder that, given the constraints, both mathematical and methodological, under which he worked, PTOLEMY was unable to "find" the law of second differences, much less the correct law. Wedded, as he understandably was, to the idea that refraction and reflection are fundamentally equivalent and, therefore, that refraction, like reflection, should be governed by a constant angular proportionality, PTOLEMY was destined to fail on at least two counts. First, it would have been little short of miraculous if, fixated upon the actual angles of incidence and refraction, he had ever stumbled upon the proper sine relation between them. Not only is this relation not actually in the phenomena as he recorded them; it was not actually in the phenomenon as he conceived it.<sup>37</sup> Second, by attempting to establish a definite relation between those angles according to the ideal  $i:r$  format, he set himself an unimaginably difficult task. Even the simple pattern of change in  $r$  that he imposed by streamlining his "observations" was too complex to be effectively expressed in terms of Euclidean proportions. But the full measure of PTOLEMY'S failure is revealed by the inadequacy of the law that he actually did propose. Unable to establish a definite general relation between  $i$  and  $r$ , he chose the only valid alternative he had: to establish an indefinite general relation between them. Thus, at the end of his experimental analysis of refraction, he concluded that, if the ray passes from a rarer to a denser medium, and if  $i > i_1$ , then  $i:r > i_1:r_1$ , which means that, instead of a constant proportionality, there is a continuous *disproportionality* between  $i$  and  $r$ .<sup>38</sup>

Suffice it to say, this disproportionality statement is both true and trivial; and its triviality has caused it to be, if not entirely ignored, at best passed over with the bland assurance that no scientist of PTOLEMY'S caliber would have seriously intended it as a law.<sup>39</sup> Perhaps not, but it was the only law he could have legitimately proposed under the circumstances. It follows logically from his experimental results, it conforms to his basic model, it is absolutely general, and most revealing of all, it takes expression in the sought-after  $i:r$  form. In fact, its very triviality and insufficiency make this disproportionality law critically important as an index of precisely what PTOLEMY was looking for in refraction and the severity of the constraints under which he carried out the search. For in the final

<sup>36</sup> This final constant relation would then be related to 10:8, 10:7, or 10:9.5, depending on which case of refraction is under consideration. Thus, for refraction from air to water, the full statement of proportionality would be  $i:\left(r + \frac{n^2 - n}{2} d_2\right) :: 10:8$ .

<sup>37</sup> In other words, because PTOLEMY quite naturally conceived of refraction as a function of the angles, he was not logically bound to take the next step of abstraction: to conceive of it as a function of a *function* (i.e., the sine) of the angles.

<sup>38</sup> This is what PTOLEMY'S final summary in *Optics* V, 34, p. 245 boils down to.

<sup>39</sup> For example, it is improbable that, without having completely ignored this claim of disproportionality in V, 34, SARTON could have maintained that "Ptolemy found that the angles of incidence and refraction are proportional . . ." (*Introduction*, vol. 1, p. 274). As is clear from his tabulations, PTOLEMY found no such thing.

analysis, the form of that law bespeaks the singlemindedness of his effort to save refraction in terms of direct angular proportionality, and its futility is a perfect gauge of the futility of that effort.

### Conclusion

Why did PTOLEMY fail to find the sine-law? Or, to rephrase the question, why did he fail to save the appearances of refraction? The answer is that it was not so much PTOLEMY who failed to save the appearances as "saving the appearances" that failed him by leading him at the outset to treat the domain of optics as a particular kind of anomaly-structure. Accordingly, he was bent on resolving that structure into a more or less rigid system of three basic anomaly-types, depending upon whether the visual ray was unbroken, reflected, or refracted. The first two of these anomaly-types, as well as their respective salvations, had been effectively reduced to equivalency through the least-lines principle, and the last two (*i.e.*, reflection and refraction) displayed a number of crucial affinities. Therefore, since the phenomena or appearances of reflection and refraction seemed to be essentially equivalent, it stood to reason that their respective salvations would be too. The obvious conclusion for PTOLEMY to draw, then, was that in refraction, as in reflection,  $i$  and  $r$  would be constantly proportional, although in the case of refraction they would be unequal.

If indeed PTOLEMY did draw this conclusion, then even before he undertook his refraction experiments, he had a very definite idea of what to expect. Yet his observations failed to bear it out. Instead of increasing uniformly with  $i$ , as he had anticipated,  $r$  seemed to increase at a steadily decreasing rate. Since there was a clear anomaly between expected  $r$  ( $R$ ) and observed  $r$ , PTOLEMY had to face the problem of rationalizing it. Guided by the assumption that Nature acts as efficiently, and thus as simply, as possible (the Principle of Natural Economy), he did so in the simplest way possible, by reducing the anomaly to two fundamental components: namely, a uniform, ideal progression of  $R$  according to constant increments, offset by a constant regression in those increments. The beauty of this solution is that, with only the subtlest violence to the raw data, it fits the "observations" perfectly.

Having arrived with no small ingenuity at a rational and simple pattern of variation between  $R$  and  $r$ , PTOLEMY then had to find a way of reconciling  $i$  and  $r$  according to it. The easiest way of accomplishing this was through the mediation of  $R$ , so that the apparently inconstant relation between  $i$  and  $r$  could be resolved into two constant sub-relations between  $i$  and  $R$  and between  $R$  and  $r$ . Easy enough in theory, this resolution proved to be virtually impossible in practice, primarily because of the limitations of Euclidean proportionality theory. The major stumbling-block would have been the problem of relating the expression

$R = r + \frac{n^2 - n}{2} d_2$  directly to  $i$ , which would have barred him from any effective

statement of law in the ideal  $i : r :: i_1 : r_1$  form. Hence, as a result of his understandable fixation on the direct  $i : r$  relation, PTOLEMY was not only incapable of "finding" the sine-law; he was forced *faute de mieux* into the weak generalization that  $i$  is invariably disproportional to  $r$ .

Ironically enough, it was precisely because PTOLEMY'S approach to refraction was so eminently sensible and logical that he strayed so far from the sine-law. Actually, he took a fatally wrong turn at the beginning, when he assumed (quite correctly) that refraction is a special case of reflection and, therefore (quite incorrectly), that its salvation must lie in a direct relation between *angles i* and *r*. Once he had batted onto that idea, he pursued it with logical singlemindedness along a path of impeccable reasoning that led inexorably to failure. As it turns out, the anomaly of refraction was not just more complex than he had originally thought; it was more complex than he could have imagined.

At a somewhat superficial level, then, PTOLEMY'S failure can be imputed to his procedure, which misled him into focusing too narrowly upon direct angular relationships. But the real source of failure can be traced even deeper, to the basic assumption upon which that procedure rested: that the visual act is ultimately intelligible in terms of simple, fixed spatial relationships between viewer and object. As the final arbiter of these relationships, the visual ray was therefore conceived of as a fundamentally spatial entity—a real line through which the geometrical reality behind the visible appearances could be immediately construed. And the rectilinearity of that line was presumed to be a function of its absolute spatial brevity.

We have already seen how successfully the ray-as-least-distance was employed in the salvation not only of direct vision but, far more important, of reflection. HERO'S demonstration of the contingency of the equal-angles law upon the Principle of Least Lines is a clear testament to that success. In the case of refraction, though, the same sort of analysis will not work. The sine-law simply cannot be established on the basis of least distances but, as FERMAT eventually showed, must be grounded upon least *times*. In other words, if it is to save refraction, the ray must be understood to represent a temporal, not a spatial, path.<sup>40</sup> The inadequacy of PTOLEMY'S analysis of refraction was therefore due to the inadequacy of his ray-concept. Like the extremal principle (*i.e.*, the least-lines principle) upon which it pivoted, it was too limited to be extended to the salvation of *dioptrics*.<sup>41</sup>

Thus PTOLEMY'S failure was at bottom systematic, a function less of *his* limitations than of the inherent limitations of the conceptual framework within

---

<sup>40</sup> Both FERMAT'S proof and DESCARTES' earlier "demonstration" of the sine-law depend upon the mathematical ray's representing not distance but velocity (*cf.* A. I. SABRA, *Theories of Light from Descartes to Newton* [London: Oldbourne, 1967]). Consequently, the ray as they finally conceived it represented not a spatial but a temporal (or spatio-temporal) entity. Moreover, by grounding his ray-analysis on least times, FERMAT established that the perfect spatial brevity of the ray was a function of its even more fundamental temporal brevity.

<sup>41</sup> The basic flaw in PTOLEMY'S refraction-analysis consists in the fact that the terms of his analysis are too concrete and specific. The spatial brevity that he supposed to be the fundamental governing principle of visual radiation is actually a function of a more profound temporal brevity. Likewise, the angular relationships that he thought governed refraction are actually functions of a more profound sine relationship. Thus, PTOLEMY'S failure overall was due to his inability to conceive the phenomena more abstractly, to transcend the limitations of the simple spatial intuitionism that dictated his scientific approach.

which he operated. The heart of the problem lay in the cardinal supposition that underpinned the whole appearances-saving enterprise: that the true and final measure of any spatio-temporal process—be it visual/visible radiation or celestial motion—is its spatial extent. Accordingly, the visible appearances and their salvation were conceived to be essentially spatial and at best only accidentally temporal.<sup>42</sup> As a consequence, while the mathematical lines used in the classical and early modern analysis of optics are identical, what they represent are decidedly not. It is this crucial difference between the ray, understood as a real line in space, and the ray, understood as an imaginary line in time, that separates classical Euclidean-Ptolemaic from early modern mathematical optics. In short, the means of discovery available to PTOLEMY were not the same as those available to his early seventeenth-century successors, because his conceptual grasp of the phenomena was radically different from, and far less abstract than, theirs.

History Department  
University of California  
Riverside

*(Received August 4, 1981)*

---

<sup>42</sup> Thus, “saving the appearances” in general was predicated on the same spatial intuitionism that vitiated PTOLEMY’S analysis of refraction. In both classical optics and classical astronomy, this led to a fundamentally static approach: the appearances were assumed to be frozen in space and, in a sense, frozen out of time. Change was thus seen as a basically spatial rather than temporal phenomenon. It was this failure to incorporate time effectively into the analysis of change that constituted the fatal flaw of the appearances-saving endeavor, and as far as mathematical optics is concerned, that flaw persisted until the static approach of EUCLID and PTOLEMY gave way to the kinetic approach of DESCARTES and FERMAT.