Comment on the Origin of the Equant papers by Evans, Swerdlow, and Jones

Dennis W Duke, Florida State University

A 1984 paper by Evans[1], a 1979 paper by Swerdlow (published recently in revised form)[2], and a forthcoming paper by Jones[3] describe possible scenarios and motivations for the discovery of the equant. The papers differ in details, but the briefest outline of their arguments is that whoever discovered the equant

1. determined the apparent value of 2*e*, the distance between the Earth and the center of uniform motion around the zodiac, corresponding to the zodiacal anomaly. One can use a trio of oppositions, some other procedure based on oppositions (e.g. Evans[4]), or in principle, any synodic phenomenon, but in practice oppositions would be the clear observable of choice.
2. determined the apparent value of e', the distance between the earth and the center of the planet's deferent, by looking at some observable near both apogee and perigee: Evans and Swerdlow use the width of retrograde arcs, Jones uses the time interval between longitude passings at the longitude of the opposition. For both kinds of observables, the pattern of variation with zodiacal longitude is found to be grossly incompatible with the predictions of a simple eccentric model. Indeed, all three find that the eccentricity e' is about half the value of 2*e* determined in the first step.
3. reconciled these different apparent eccentricities with the invention of the equant, which by construction places the center of uniform motion twice as far from Earth as the center of the planet's deferent.

The scenarios of Evans, Swerdlow and Jones are all perfectly possible, and consistent with what we know of the early practice of ancient astronomy. However, there is yet another path that might have led to the discovery of the equant, and it is in some ways both simpler and more general than those just mentioned. Indeed, let us imagine that an early analyst, not yet knowing about the equant, of course, is using an ordinary eccentric model to try and understand the spacings in time and longitude of the oppositions of Mars with the mean Sun. The analyst will assemble the required empirical data by making a series of timed longitude measurements of the planet during retrograde, and comparing with the position of the mean Sun predicted by his solar model, determine the time and longitude of opposition. He would then execute a trio analysis of the oppositions, which would reveal the eccentricity 2*e*, the longitude *A* of the apsidal line, and the mean longitudes of the epicycle center at the time of each opposition. For Mars, he would find that 2*e* is some value near 12 and that *A* is some value near 115°. If he analyzed more than one trio of oppositions, he would find values scattered around these values, and would soon gain confidence that his analysis was trustworthy so far.

Having determined the parameters required to model the oppositions, the analyst might well decide to proceed with the determination of the epicycle radius *r*. He would certainly need the value of *r* to, for example, study retrograde arcs. To find *r* he can simply take any one of the timed planetary longitudes he used to determine one of the oppositions, or any other measurement of the planet's position near the opposition that he might have handy, and a short and simple calculation[5] will give him *r*. Ptolemy, in fact, does precisely this, once for each superior planet, in *Almagest* 10.8, 11.2, and 11.6. Now while Ptolemy determines *r* using only the third opposition in his trios for each planet, it would certainly be natural for the analyst who didn't already know the answer to estimate *r* not just once but *three* times, using *each* of the oppositions in his trios, and if he had indeed already analyzed a few more trios, he could get three more values of *r* from each set. If he did so, and it is hard to believe he wouldn't, the results he would find, for Mars, would reveal a striking problem: the values determined for *r* would very likely be significantly different, ranging from about 35 to over 43. This would, of course, be an intolerable situation, conflicting as it does with the standard Greek cosmological picture of spheres of the planets, and our analyst would have no choice but to look deeper into the problem.

It is likely that his next step would be to ask if there is some systematic pattern in the variation of the *r* values. It wouldn't take many values of *r* to reveal that the variation is in fact a simple function of zodiacal longitude, varying in a fashion that we would call sinusoidal, with a maximum near Mars' apogee (about 115°), and a minimum near perigee (about 295°). The data in the table show what Ptolemy, for example, would have gotten if he had analyzed all the oppositions between 124 AD and 141 AD.

year	*longitude*	*r*
124	271	35.5
128	42	40.1
133	116	43.4
126	348	36.4
130	82	42.7
135	150	42.7
128	42	40.1
133	116	43.4
137	190	39.7
130	82	42.9
135	150	42.3
139	243	36.0
133	116	43.4
137	190	39.7
141	320	35.6

Then at some point he might well think along the following line: I have arrived at this dilemma by assuming that my distance from apogee is about $R + 2e = 72$ and my distance from perigee is about $R - 2e = 48$. Is there any way I can check if this is correct? In fact, what eccentricity would I find if I could somehow *require* that I get the same value of *r* at both apogee and perigee? This calculation is a straightforward variation of the calculation used to get the r values in the first place. If he looks through his database, he will find that the opposition in 133 AD occurred very near apogee, and the oppositions of 124 and 141 AD occurred within about 20° – 25° on each side of perigee. Now the final equation that determines r in the usual analysis is the solution of the triangle defined by the Earth – planet – epicycle center, which gives

$$r = \rho \frac{\sin p}{\sin(\alpha_v + p)}$$

where ρ is the distance from the Earth to the epicycle center, p is the equation of anomaly, *i.e.* the angle subtended by the epicycle as seen from the Earth, and α_v is the true anomaly, i.e. $\alpha_v = \alpha + q$, where α is the mean anomaly and q is the equation of center. Now near apogee ρ is close to $R + e$, and near perigee ρ is close to $R - e$. Furthermore, near both apogee and perigee the equation of center q, is quite small, and of course $q = 0$ exactly at apogee and perigee. Thus let us assume that the observations of 124, 133, and 141 are indeed on the apsidal line, and that the fact that they are not will not cause a noticeable problem (a fact easily verified *a posteriori* by our analyst). Then the requirement that r be the same at both apogee and perigee is simply

$$(R+e)\frac{\sin p_{\max}}{\sin(\alpha_{\max} + p_{\max})} = (R-e)\frac{\sin p_{\min}}{\sin(\alpha_{\min} + p_{\min})}.$$

Letting $a_{\max} = \dfrac{\sin(\alpha_{\max} + p_{\max})}{\sin p_{\max}}$ and $a_{\min} = \dfrac{\sin(\alpha_{\min} + p_{\min})}{\sin p_{\min}}$, e is given simply by

$$e = R\frac{a_{\max} - a_{\min}}{a_{\max} + a_{\min}}.$$

Using the values calculated from the circumstances of the observations of 124 and 133, one finds $a_{max} = 1.656$, $a_{min} = 1.363$, and hence $e = 5.8$. Using instead the observation in 141 gives $a_{min} = 1.372$ and hence $e = 5.6$. Both values of e are then seen to be about half the value of $2e$ found from analyzing the zodiacal anomaly. And in any event, if the analyst decided he needed observations very close to the apsidal line, he can get them about once a year for Mars. Everything would then be in place for the invention of the equant. Probably the principal remaining difficulty would be to figure out how to analyze a trio of oppositions in the new equant model. The iterative solution documented in the *Almagest* is by any measure a piece of brilliant mathematical analysis.

So the simple scenario of discovery might well have been the analysis of a few trios of oppositions, and a notice that the radius of the epicycle obtained from a short analysis of the same data gives values of r that vary with zodiacal position. Besides simplicity, this explanation of a path to the discovery of the equant has the following virtues:

1. the empirical data required for the analysis of r is already in the database of the analyst, although he certainly could make additional observations if he wanted.

2. the computational method used to determine *r* is explicitly attested in the *Almagest*. It is also simple enough that multiple determinations of *r* would not be an undue burden on the analyst.
3. it is hard to imagine that any analyst would fail to notice the variation in *r* obtained by using a variety of oppositions at positions around the zodiac.
4. the method is consistent with Ptolemy's description that "…in the case of each of these planets [Mars, Jupiter, and Saturn], speaking in terms of a rather rough method, the eccentricity that is found by means of the greatest difference caused by the zodiacal anomaly proves to be approximately double the eccentricity derived from the magnitude of the retrogradations [of the planet] around the greatest and least distances of the epicycle".[6] While that language certainly includes the retrograde arcs analyzed by Swerdlow and Evans, it could also include the longitude passing suggested by Jones and, more generally, the observations near opposition suggested in this paper.
5. this method of exposing and solving the problem for the outer planets corresponds closely to the method Ptolemy describes in *Almagest* 10.2 to justify the equant for Venus, namely the eccentricity of the deferent is that required to keep the apparent size of the epicycle the same at apogee and perigee, and it is about half that obtained by analyzing greatest elongations near quadrature.
6. While the variations in the size of *r* for Jupiter (about 11.4 – 11.9) and Saturn (about 5.8 – 6.8) are not as large as for Mars, they are certainly large enough to serve as confirmation of the idea so strikingly motivated by the Mars data.

This investigation, like all the others, in principle leaves open the question of *when* and *who* first discovered the problem from empirical data, proved that a bisection of the eccentricity was needed, and invented the equant to implement that solution. Most people have, of course, credited all that to Ptolemy himself, but the main evidence is simply that the equant is never mentioned in any source older than the *Almagest*. Ptolemy's own words in the *Almagest* are, however, somewhat ambiguous on the subject. While he certainly does not credit anyone else for the discovery, which he did, for example, in the case of Hipparchus and the solar model of *Almagest* 3, he also does not explicitly take credit for himself.

We are then left with the question of timing. Does what we know of pre-*Almagest* astronomy support the possibility of an earlier discovery? We have seen that the ingredients necessary for progress down the path outlined in this paper are (a) a reasonably reliable solar model, (b) an ability to measure a sequence of timed planetary longitudes during retrograde in order to determine opposition, and (c) the mathematics to do an eccentric model trio analysis. We know for certain that part (a) was satisfied no later than Hipparchus, for it is multiply attested. Part (b) is also very likely, since how else could theorists have made even the mediocre progress that Ptolemy disdains in his discussion of planetary model history in *Almagest* 9.2? It is almost as certain that the description of Hipparchus' analysis of lunar trios in *Almagest* 4.11 means that part (c) was satisfied also no later than Hipparchus. Thus, based solely on timing, there is no apparent reason that the bisection of the equant could not have been discovered at the time of Hipparchus, and as far as we know, perhaps even a bit earlier.

But there is, in fact, another mention of the bisected equant in ancient astronomy[7], and it might bear on the question of when the equant was invented. It so happens that the bisected equant is clearly attested, albeit in a somewhat disguised form, in the planetary models of ancient Indian astronomy.[8] While the oldest texts we have that contain full planetary models are thought to originate in India sometime in the period 400-500 AD, it is also generally thought that the underlying astronomy in these texts is Greco-Roman in origin, and from a time that predates the *Almagest*.[9] The reason for this dating is that the astronomy found in the texts is, on the whole, considerably less developed than what we find in the *Almagest*. Therefore it might be hoped that the bisected equant, by any measure one of the most sophisticated elements in the *Almagest*, being found among this heap of more primitive astronomy will motivate a careful new analysis of ancient astronomy, perhaps from a new perspective.

ACKNOWLEDGEMENT

I would like to thank Noel Swerdlow for a critical reading of the manuscript and for several suggestions that led, hopefully, to a clearer presentation of the ideas.

REFERENCES

---

[1] James Evans, "On the function and probable origin of Ptolemy's equant", *American journal of physics,* 52(1984), 1080-9.
[2] Noel Swerdlow, "The empirical foundations of Ptolemy's planetary theory", *Journal for the history of astronomy*, 35 (2004), 249-71.
[3] Alexander Jones, "A Route to the ancient discovery of nonumiform planetary motion", *Journal for the history of astronomy*, 35 (2004).
[4] James Evans, *The history and practice of ancient astronomy* (New York, 1998), p 364.
[5]The details of the calculation may be found in Olaf Pedersen, *A Survey of the Almagest*, (Odense, 1974), 285-6, or Otto Neugebauer, *A History of Ancient Mathematical Astronomy,* (New York, 1975), 179-80.
[6] this translation is by Alexander Jones, *op.cit.* (ref. 3).
[7] B. L. van der Waerden, "Ausgleichspunkt, 'methode der perser', und indische planetenrechnung", *Archive for history of exact sciences*, 1 (1961), 107-121 made the original suggestion and offered a justification based on a power series analysis. more recently Dennis Duke, "The equant in India: the basis of ancient Indian planetary models" (2004, submitted), has presented detailed numerical comparisons that remove any doubt that in fact the bisected equant underlies the Indian planetary models.
[8] D. Pingree, "The *Paitamahasiddhanta* of the *Visnudharmottapurana*", *Brahmavidya*, xxxi-xxxii (1967-68), 472-510; K. S. Shukla, *Aryabhatiya of Aryabhata* (1976); B. Chatterjee, *The Khandakhadyaka of Brahmagupta* (1972); K. S. Shukla, *Mahabhaskariya of Bhaskara I* (1960).

[9] D. Pingree, “On the Greek Origin of the Indian Planetary Model Employing a Double Epicycle”, *Journal for the history of astronomy*, ii (1971), 80-85; D. Pingree, “The Recovery of Early Greek astronomy from India”, *Journal for the history of astronomy*, vii (1976), 109-123; D. Pingree, “History of Mathematical astronomy in India”, *Dictionary of Scientific Biography*, 15 (1978), 533-633.