9. SOME ASPECTS OF MODEL FITTING

9.1 INTRODUCTION

The previous chapter concerned Mayer's work of *model fitting*, as we would call it nowadays. Mayer had to deal with errors in his model, in its coefficients, and in the observations that he employed to fit it. In fact, Mayer's work abundantly shows his ambition to take control over errors, and to increase accuracy and precision: be it in cartography, the mapping of the moon's surface, or the lunar orbit, or the consistent description of colours. It is evident in his design of instruments, in the 'repeating principle' that he introduced in angle measurements, in the *Mappa Critica*; it is the unifying theme in his work, as Forbes has already pointed out. 'The science of practical errors is so far not yet sufficiently developed,' wrote Mayer when setting out on an investigation into the limits of the human visual acuity under various light conditions, in order to ascertain the accuracy of angular measurements.[1]

We continue now with other aspects of Mayer's work where his commitment to getting the best possible result out of the available data shows through. We will gain some insight into the statistical aspects of his work. Mayer developed ad-hoc procedures in each case; the statistical tools of today had not yet been developed.

First, we investigate the quality of the observations at Mayer's disposal, particularly those which he used in the spreadsheets. This leads us to the realization that the result of his fit is almost optimal, but that the precision to which he worked was not matched by the accuracy of the data. Next, we briefly discuss a memoir of Mayer's in which he proposes a mathematical model of world temperature distribution. I argue that this memoir bears testimony to Mayer's trust in his spreadsheet tool. Then we turn to older work of Mayer's. We investigate his use of the arithmetic mean and his awareness of the cancelling property of random errors. Section 9.5 details another of Mayer's attempts to fit a model to observations: he fits a model of lunar libration to data via a redundant system of linear equations. Section 9.6 is devoted to Mayer's further use of redundant systems of linear equations and their role in the success of his lunar tables.

9.2 THE QUALITY OF OBSERVATIONS

As we have seen (particularly in the previous chapter), Mayer went to great lengths in fitting the lunar tables to observations. An interesting question concerns the quality of the observational data on which the fit was based, more specifically: how

1 *'Scientia errorum practicorum nondum satis hactenus exculta'* [Mayer, 1755, p. 120].

much random error is present in the data, and to what extent is Mayer aware of that. In order to investigate these questions, I sampled the position calculations in Cod. $\mu_{41}^{\sharp}$ between fol. 31r and 72v, which are all related to occultations of Aldebaran. Mayer used almost all of these in his spreadsheets, rejecting only those that seemed to yield very unlikely results to him. These occultations allow us to obtain a unique impression of the quality of data at Mayer's disposal, as I will explain next. The result enables us to see Mayer's fitted tables in a new quantitative perspective.

An occultation of a star by the moon provides two sharply defined observable phenomena: the disappearance of a star behind the disc of the moon (called its *immersion*), and its subsequent reappearance (or *emersion*). The observation of an occultation consists mainly of the recording of the (local) times of these phenomena. When due consideration is given to the lunar parallax and diameter, and to the location of the point of contact on the lunar disc relative to its direction of motion, the observations provide very accurate positions of the moon relative to the star. This property makes occultations particularly suitable for Mayer's purpose.

The temporal separation between the two phenomena can never be much longer than an hour, because in that time interval the moon appears to move approximately its own diameter relative to the stars. Such a time-span is very short compared to the periods of the lunar inequAlities. Therefore the errors of the periodic equAtions in the lunar tables before and after an occultation are equal for all practical purposes. Mayer already had the mean motions approximately right, so these too will produce equal errors before and after the occultation. In conclusion, the predicted lunar position will differ from its true position just as much at immersion as at the emersion. This property makes *pairs* of occultations particularly suitable for our purpose as stated in the introduction to this section. Of course, they are convenient for Mayer's purpose, too.

Mayer had a collection of occultation observations from his century and the previous one, by several observers, and from various locations. These were the most important data that he applied in the spreadsheets to improve the coefficients of his tables. The position calculations that Mayer made to that end, or rather the immersion-emersion pairs among them, enable us to pursue our current aim. I refer to the example calculation in figure 8.1 on page 143, showing an Aldebaran occultation of 2 January 1738 observed from Paris. The calculation occupies two facing folios, the left side for the immersion, and the right side for its emersion.

In that example, Mayer computes the lunar position from the `kil` tables, following the general scheme set out in section 4.1. This computation occupies approximately the top one third of both folios. Below that, Mayer reduces the observations, allowing for aberration and parallax (but not for nutation).[2] Apparently he has two

2 The fast change of parallax in right ascension provides the reason why it is not possible to average the times of immersion and emersion and make a position calculation for the middle of the occultation only, thus saving half the work. The effects of aberration and nutation can add about $20''$ and $17''$ respectively to lunar longitude; however, since they are slow-changing, they have no impact on the sequel of this paragraph. (Aberration is an apparent deflection of light rays, resulting from the finite speeds of light and of the earth.)

different observations of immersion and also two of emersion.[3] Consequently he derives four errors, i.e., differences between computed and observed positions of the moon: for immersion, they are $-33''$ and $-21''$, and for emersion, he finds $-20''$ and $-28''$. He concludes that the average error is $-25''$.[4]

This averaged result is a measure of agreement between observations and theory for the specified instance, which of course is precisely why Mayer had made these calculations. But to investigate the data quality, our attention is now drawn to the dispersion around the average: it shows us the relative quality of two different observers, each observing one immersion and one emersion. Observer number one obtained a dispersion of $|33'' - 20''| = 13''$, but number two had a sharper result of $|21'' - 28''| = 7''$. These dispersions are independent of the quality of the lunar tables, because the error of the tables is the same at the beginning and end of the occultations. But they do reflect errors in the observations and their reduction.

In this way I investigated 34 position calculations of occultations where I could distinguish both immersion and emersion, skipping those that Mayer rejected. The data spanned the period from 1680 to 1750. I found a dispersion of $10''$ or less in 20 cases, $10''$ to $20''$ in 5 cases, and the remaining 9 cases showed a dispersion between $30''$ and $69''$.[5] It is illustrative to compare these numbers to the standard deviations that I computed for Mayer's spreadsheet results as summarized in appendix B. Mayer accomplished to bring the standard deviation down to $16''$, very close to the least-squares fit of $14.5''$. Our current investigation shows that the quality of his fitted tables matches the quality of his data. This means that Mayer could not have achieved a significantly better result than he did.

3 In Cod. μ_{12}, where Mayer collected his data, there are two references to this occultation: on fol. 63v, copied from the Mémoires for 1739, and on fol. 80v referring to correspondence between Lowitz and P.C. Maire. He attributed the second observation to Lemonnier.

4 The arithmetic mean of all four results is $-25\frac{1}{2}''$.

5 These data lead to the following observation. Taking half of the stated dispersion as the observation error (which is admittedly a best-case scenario), we see that Mayer's own estimate of $5''$ to $10''$ for these observations is slightly optimistic ('We have a considerable number of observed occultations of Aldebaran; out of these I have calculated the positions of the Moon with the help of the parallax, so that there can be no error of $5''$ or $10''$ in them' [Forbes, 1971a, p. 81]). The argument that the errors in the lunar tables are equal before and after the occultation, can here be repeated for the clock error. The quality of the observations, which I have just related to the difference of errors before and after the occultation, depends mostly on the ability of the observer to trap the exact instance of the star's disappearance and reappearance at the limb of the moon: a longitude error of $10''$ is equivalent to a clock error of about 20 seconds. Before immersion, the observer can see the star approaching the limb, whereas the emergence is sudden; this makes the immersion (in principle) easier to observe. Other factors that play a role include the observer's experience, the meteorological circumstances, and whether the moon is waxing or waning (i.e., which of the moon's limbs is illuminated).

Incidentally, the quality of the data at Mayer's disposal (as here investigated from pairs of observations) matches the quality that Tycho Brahe achieved in his determination of the longitude of α Arietis, after the latter had combined his data in pairs to eliminate certain systematic errors. This conclusion shows clearly the unprecedented accuracy attained by Tycho, a conclusion that is only slightly moderated when it is taken into account that Tycho built his data set personally and for the specific purpose of fixing the longitude of α Arietis as a reference star. For details on Tycho's determination, see [Hald, 1990, pp. 145–6] and [Plackett, 1958].

Page after page, occultation after occultation, Mayer must have noted the dispersion in individual observations. Seeing such a dispersion did not restrain him from filling in his spreadsheets to a precision of half a second. For us, this is hard to understand. 'Unwarranted number of significant digits,' we tend to remark. We encountered another example of this when we studied the relation between the single-step and multistep equAtions in *Theoria Lunae*, on page 128. Donald Sadler observed the same phenomenon in the *Nautical Almanac* and in Nevil Maskelyne's procedures for the computation of (geographical) longitude by lunar distances, and the accompanying tables. With one significant digit less (i.e., replacing minutes and seconds by minutes and tenths of minutes), the resulting longitude would have suffered inappreciably, but the burden on the human computer would have decreased considerably, mainly because a number of corrections could be skipped on account of their small size.[6] Certainly numerous other examples exist of this phenomenon of over-precision in the 18th century. It seems to me that an attitude towards data prevailed in which a *lack of accuracy* (such as the dispersion in the input data at Mayer's disposal) was accepted, by many even recognized as partly inevitable; the self-cancelling property of random errors was recognized by many. But at the same time we see both Maskelyne and Mayer on their guard against any *loss of precision* without realizing the futile part of their effort in view of the limited accuracy of the data.

9.3 WORLD TEMPERATURE DISTRIBUTION

In the memoir *A More Accurate Definition of the Variations of a Thermometer*,[7] Mayer expounds a mathematical model for computing the temperature of places on earth, as a function of various geographical data. He specifies a formula for the average annual temperature of a place as a function of the latitude. Then he provides several refinements in the form of: a correction term for the altitude, a term for annual (seasonal) variation, and a term for diurnal variation. Mayer asserts that particular attention must be given to the period, phase, and amplitude of the last two variations. These are governed by certain coefficients in the formulae, whose values are quite easily obtained with the help of observations. Mayer gives only a few simplified examples, but no thorough treatment, of the determination of those coefficients. He does not claim any predictive power for his model. Rather, his main goal is to propose a method to investigate the actual temperature data: 'I think that it is impossible to define the causes and quantities of the remaining, more involved, variations, unless the effects of the different causes are analysed in the way which I have roughly sketched out here.'[8]

6 [Sadler, 1977, pp. 115–119].

7 Translated and republished in [Forbes, 1971c], original Latin text posthumously published in [Mayer, 1775].

8 [Forbes, 1971c, p. 61].

Forbes and Delambre have commented that the interest of this memoir by Mayer lies more in the method than in the results.[9] Mayer himself had put forward that astronomers have good methods to investigate the movements of the luminaries, and that meteorologists have something to learn from astronomers:

> Therefore, at this point, we may transfer the example of that astronomical method to variations in the atmosphere, applying it in particular to the ratio of heat and cold: it will thereby be possible to show how I deem that meteorological observations ought to be treated, so that richer fruits can be expected from them.[10]

The relation between Mayer's memoir and his interest in correcting raw astronomical observations for atmospheric refraction, which depends on the temperature, has already been noted by Forbes.[11] Forbes summarizes Mayer's proposed method (in the *Thermometer* treatise) as 'isolating a major periodicity, examining the residuals for a second-order periodicity, etc., until the observations had been analysed into a series of independent periodic functions each characterized by a mean value and its variation about the mean'[12] and concludes that the relationship to Mayer's lunar investigations is obvious, aiming at the analogy of periodic fluctuations, and the customary way among astronomers to investigate orbital motions. However, a major difference between the models, not pointed out by Forbes, is that the lunar theory was already able to supply *a priori* approximate values for the amplitudes of the anticipated periodicities, a feat which was (in Mayer's era certainly) unmatched by the theory of heat.

It should also be noted that Mayer had such a strong confidence in his astronomical perturbation analysis techniques, that he did not hesitate to export them to other disciplines. Certainly Mayer's approach in the *Thermometer* treatise bears witness to both his interest and his confidence in fitting models with many periodic terms. Significantly, he read the *Temperature* memoir to the Göttingen Scientific Society on the 13 September 1755, when he had already developed the spreadsheet tool and when his *Theoria Lunae* was just about finished. It is, in short, an attempt to export the tools of his astronomical work to another domain. With some imagination, we can picture Mayer studying spreadsheets of temperature data.

9.4 AVERAGING AND CANCELLING

The rest of this chapter is concerned with aspects of data use and model fitting present in Mayer's work before 1751, i.e., when he was still working with the Homann heirs in Nuremberg. Currently, we will investigate two related aspects of working with data: the fact that averaging over several data usually leads to a better result, and the property of random errors to cancel each other. A probabilistic

9 [Forbes, 1971c, p. 21], [Delambre, 1827, p. 447].
10 [Forbes, 1971c, p. 54].
11 [Forbes, 1980, pp. 178–181].
12 [Forbes, 1971a, pp. 15–16].

proof that the arithmetic mean of a series of observations is more reliable than an individual observation was first given in 1755, by Thomas Simpson.[13]

In a memoir *Untersuchungen über die geographische Länge und Breite der Stadt Nürnberg*, Mayer evaluated older observations made by Wurzelbau in Nuremberg in order to determine the latitude of his habitat.[14] As Forbes mentioned,[15] Mayer's intention was to warn against the use of results (in this case, the latitude of Nuremberg) without a proper investigation into their origin (here, Wurzelbau's observations). I regard Mayer's memoir as illustrative of his use of data from other astronomers, and he may even have started to work on it to show his skill.

First, Mayer gave four series of Wurzelbau's observations: two series of superior culminations of Polaris, and two series of inferior culminations. He rejected the inferior culminations because the individual observations in them deviated considerably more from the mean than those in the superior culminations. This shows that he considered the reliability of his data, and that he was prepared to keep only the most reliable ones.

Next, Mayer averaged the remaining series, as was not uncommon among astronomers, and then computed Wurzelbau's latitude from the mean altitude of Polaris at upper culmination (he also used an independent determination of the polar distance of that star, which he could otherwise have derived from the combination of the superior and inferior conjunctions).

Mayer also explored two other series of observations made by Wurzelbau by means of different instruments, of culminations and solstices, but these comprised much shorter data series. Mayer used each of those other data series for a separate computation of the latitude of Nuremberg. These several different latitude determinations were apparently in reasonable agreement, at least after Mayer had corrected the results for errors in Wurzelbau's instruments which he had detected through his careful analysis of the observations. Two of the three results differed by less than $2''$; the third differed from both by almost $20''$. Mayer felt justified to average over all three results, and then to round off towards the 'outlier', thereby implicitly giving it a slightly larger weight. We see that he was careful to cross-check results before he relied on them.

There are certain differences between his procedure and modern ways of data handling. Mayer averages his computed latitudes without weighing them, although the data series from which the results were obtained are of unequal length and of unequal quality. So Mayer's end result is an average of several averages of unequal reliability. Instead, a modern statistician would prefer to compute latitudes from Wurzelbau's observations individually, and average only one time over the complete corpus, perhaps with weights assigned if the latitudes are of unequal reli-

13 [Plackett, 1958, p. 124].

14 It was published posthumously in [Forbes, 1972, I, pp. 33–44]. The memoir deals only with the latitude of Nuremberg, and the title concedes that the memoir remained unfinished. An article in the Göttingen *Commentarii* [Mayer, 1752], also on the latitude of Nuremberg, is mostly concerned with Mayer's own observations.

15 [Forbes, 1972, I, p. 9].

ability. Stigler points out that until the second half of the 18th century, astronomers were willing to average only among observations that were taken under comparable circumstances (same observer, same instrument, same object, etc.).[16] A procedure of taking weighted averages had already been outlined by Roger Cotes.[17] Mayer was evidently willing to average results obtained under unequal circumstances. Yet, he did not give these results unequal weights.

We now turn to the cancelling property of random errors. While in Nuremberg, Mayer had started to write a treatise on map making, *Von der Construction der Land-Karten*, which was to remain unfinished. One of its topics was a discussion of the value of Roman *itineraria* for map making. Mayer remarked that those works usually specified distances in rounded Roman miles of 1000 feet, without a fractional part, so the distances were not exact. But he added that sometimes these distances would be too large, and sometimes too small, and he recognized that generally the rounding would have no appreciable effect on the end result. Thus, Mayer showed that he understood a basic property of random errors.[18]

The property that random errors tend to cancel each other is also present in Mayer's design of two angle measuring instruments, the *repeating circle* and a modified *recipiangle*. These instruments are designed to accumulate a series of angle measurements for the observer. The repeating circle, for instance, allows the observer to measure an arc not only once but several times in succession, without intermediate readings of the instrument. The individual measurements are automatically summed on the circular scale of the instrument, and the observer has only to read the accumulated sum. After division by the number of observations, he obtains the average of his individual arc measurements. In this way several errors are averaged out, including the (fixed, and therefore systematic) errors in scale division, and collimation error. However, the observer can no longer recognize and discard outliers.[19]

Buchwald points out that Mayer (and Borda, who developed the instrument further) had in mind to reduce what we now call the systematic errors in their instruments, not the random errors caused by the observer. Buchwald goes on to show that the instrument accomplishes just the opposite: that it fails to average out in particular the errors in scale division.[20] That would be true if the scale had not been circular—but since it is, the scale divisions are necessarily correct *on average*, so that as long as the repeated measurements cover all parts of the circle, this kind of error is indeed averaged out. Also in his design of another instrument, a new astrolabe for surveying, Mayer explicitly tried to control random operator error.[21]

16 [Stigler, 1986, p. 30].
17 See [Gowing, 1983, pp. 107–9].
18 [Forbes, 1972, I, p. 49].
19 The recipiangle is an other angle measuring instrument, used in surveying, which will not be further discussed here. See [Forbes, 1980, pp. 153–154, 158–164].
20 [Buchwald, 2006, p. 568].
21 [Forbes, 1971b, p. 114].

9.5 LIBRATION: A CASE OF MODEL FITTING

An investigation of Mayer's work and model fitting has to take Mayer's lucidly written tract on the rotation of the moon *Abhandlung über die Umwalzung des Monds*[22] into account. Whereas Mayer himself was perhaps more interested in the direct results that he had obtained there, his work attracted the attention of many for his original approach to the much more general topic of the combination of observations, which in his particular case resulted in an overdetermined system of equations for three parameters.

This work of Mayer's, which we will now discuss, was reviewed by Lalande[23] and generalized into a method by Laplace; the latter's generalization was widely known as 'Mayer's method' and as such it was used until the first half of the nineteenth century, among others by Laplace's assistant Delambre.[24] It is quite likely that Mayer obtained his ideas from a memoir by Euler.[25] Later researchers have assigned at least two other names to it: *Method of Equations of Condition*, and *Method of Averages*.[26] The first is rather nondescript, and that is why I prefer the second name—it catches the essence of the procedure, even though it is factually wrong because Mayer did not take averages. Neither of these names were used by Mayer; indeed we may well ask whether he regarded the method really as a *method* or rather as an ad-hoc procedure. He seems to have made little or no use of it in other places, and in his hands it is less general than is perhaps apparent at first sight. Although I prefer not to regard it as a method, I will adhere to the conventional naming of method of averages.

After an introduction to the topic of Mayer's tract, and a technical exposition of the geometry of the problem he endeavoured to resolve, we discuss his method of averages in section 9.5.2. Quite naturally, the question arises whether Mayer used this same method to adjust the coefficients of the tables of lunar motion.[27] Above, I showed that he had a different procedure for that purpose around 1752 and later, but the two methods are not mutually exclusive. To find a more definitive answer, I looked specifically for places where Mayer uses similar systems of equations. Some of these I discuss in section 9.6.

The goal of Mayer's research on lunar libration was to improve the mapping of the visible lunar surface. This was important partly because of its relation to the determination of geographical longitudes via the timing of lunar eclipses, and

22 [Mayer, 1750a].
23 [Lalande, 1764, 1st ed., pp. 1234–43].
24 [Stigler, 1986, p. 31].
25 1st ed., [Lalande, 1764, p. 1241]; the Euler memoir is [Euler, 1749a]. Also see the final section of this chapter.
26 The first of these names is used for instance by Forbes and Wilson [Forbes and Wilson, 1995, p. 66], cited below. In a different place in the same book, Schmeidler uses the latter name, reserving the former for the equations that make up the overdetermined system [Schmeidler, 1995, p. 201].
27 See the quote from Forbes and Wilson on page 137.

partly because Mayer had set himself and the Homann cartographic office annex Cosmographical Society where he was then working, the prestigious task of producing accurate lunar globes.[28] Although Hevelius and Riccioli had mapped the visible part of the surface of the moon, and Jean-Dominique Cassini had initiated an investigation of the libration (discovered by Galilei and explained below), Mayer complained that the current state of selenography was rather poor: there was no consensus in the nomenclature of lunar features, and Cassini's research was not up to the attainable standards of accuracy.

It is well known that the moon always turns the same side of its surface towards the earth. Upon closer inspection this turns out to be only approximately true. For several reasons the moon is subject to a slight apparent wiggling, called libration. The reasons for this wiggling are as follows. First, due to the diurnal motion of the terrestrial observer, his aspect of the moon varies between moonrise and moonset. Second, the moon rotates (practically) uniformly around its axis while its velocity of revolution around the earth varies: consequently, a terrestrial observer sees sometimes a bit more of the leading half of the moon's surface, and sometimes a bit more of the trailing half. Third, the rotational axis of the moon is not perpendicular to its orbital plane. As a consequence, we look upon the moon's north pole for half a month, and on its south pole during the other half. The fourth effect is the slightly perturbed rotation of the moon due to its deviation from perfect sphericity. This effect is much smaller than the other librations.

An accurate mapping of the lunar surface can only be arrived at when these librations are taken into account. The first two librations depend mainly on the parallax, longitude, and latitude of the moon, which are more or less observable. The third libration depends on the orientation of the lunar axis of rotation. It is this orientation, and Mayer's investigations of it, that will concern us here. The fourth libration is too small to be observable for Mayer: this so-called physical libration was first hypothesized by Newton, but only in the 1840's were sufficiently accurate measurements available to prove its real existence.[29]

9.5.1 Locating the rotational axis

In section 13 of his libration tract, Mayer sets out to determine the orientation of the lunar polar (or rotational) axis. I will now first show how Mayer derives equation (9.2). Subsequently, I will show how he uses observations to determine the unknown quantities in that equation, and how he establishes an error estimate as well. Readers uninterested in the derivation may skip to equation (9.2).

28 See fn. 4 on p. 25.

29 A non-spherical moon had entered celestial mechanics in a different way, too. D'Alembert and Euler had tested, independently of each other, whether a non-spherical mass distribution of the moon could be held responsible for the missing part of the motion of the lunar apse line. They concluded that it could—but only if the moon were extremely dumb-bell shaped [Waff, 1995, p. 40].

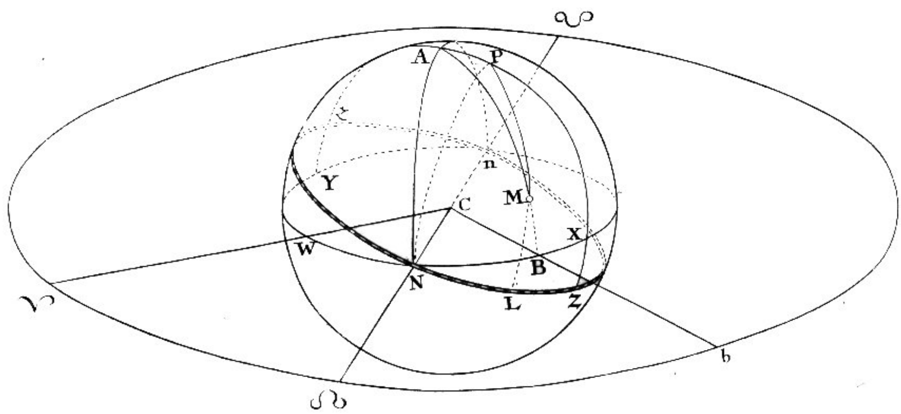

Figure 9.1: Determination of the position of the moon's polar axis. Image from the digital copy of [Mayer, 1750a] available on-line from Universitätsbibiothek Augsburg, `www.bibliothek.uni-augsburg.de`.

The geometry of the problem is illustrated in figure 9.1, which is taken from the original source. The figure represents the moon with its centre in *C*. Let the moon rotate around an axis through *CP*, with *P* its north pole. Let the equator of the moon be the great circle *NLZnz*. Draw the great circle *YWNBXn* parallel to the ecliptic; Mayer calls it the *lunar ecliptic*. Let point *A* be a pole of it.

The rotational axis of the moon is not perpendicular to the ecliptic. Therefore the planes of the lunar ecliptic and equator intersect in the line *NCn*. The points *N* and *n* on the lunar surface are called the *equinoctial points*. These points are not fixed relative to the surface of the moon; instead they both traverse the complete equator *NLZnz* during one revolution of the moon about its axis.

The circle ♈☊*b*☋ is the lunar ecliptic projected onto the celestial sphere from the centre *C*. Mayer chooses, in accord with common practice, the direction of the vernal equinox *C*♈ as the zero point of ecliptic longitude. A little less common is his use of the symbol ☊, as I will now explain. Cassini (I) had found in earlier research that the longitude of the equinoctial point *N* coincides with the longitude of the ascending node of the lunar orbit. Mayer comments that such is unlikely to be true all the time, because the motion of the orbital nodes is perturbed by the attraction of the sun while the lunar rotation is uniform.[30] Still, Mayer chooses the symbol ☊ (normally associated with the ascending node of the lunar orbit) to represent the point on the celestial sphere that corresponds with the equinoctial point *N*.

In the previous sections of his tract, Mayer had described how he had observed, between April 1748 and March 1749, 37 positions of Manilius, a distinct crater on

30 Mayer did not consider perturbations in the rotation of the moon due to its deviation from perfect spherical shape. This is the fourth form of libration mentioned in the introduction to this section. In a memoir of 1780, Lagrange concluded that the phenomenon discovered by Cassini, when taken in the mean sense, was a consequence of this physical libration [Wilson, 1995c, p. 112].

the moon not far from the equator, represented by M in the figure. Of the 37 observations, Mayer had made a selection of the 27 most appropriate ones.[31] After extensive reduction, he deduced from these 27 observations equally many values for the following quantities: Manilius' *ecliptic* polar distance $h = AM$; Manilius' ecliptic longitude $g = \Upsilon b$; and the mean longitude of the orbital ascending node k, which is *approximately* the arc $\Upsilon\Omega$ in the figure (remember that Mayer's Ω here corresponds with the equinoctial point N). Actually Mayer took the value of k from existing lunar tables for the time of observation.[32]

Mayer's goal of finding the orientation of the lunar axis entails the determination of the angle $\alpha = AP$ between the ecliptic poles and rotational poles, and the *precise* longitude $\Upsilon\Omega$ of the equinoctial point N. He represents the latter angle by $k+\theta$, with k (already determined as explained above) the longitude of the ascending node, and θ (which is not in the figure) the unknown and constant small arc separating the ascending node from the equinoctial point. As we will see, he will also have to determine the latitude of Manilius $\beta = LM$. Mayer presents evidence (which I omit here) that the orientation of the lunar axis is fixed or, in the worst case, changing so slowly that it can be regarded as fixed over a time span of a year or so.

To derive a relation between the known quantities g, h and k, and the unknown parameters α, β and θ, Mayer considers the spherical triangle APM. The cosine rule of spherical trigonometry, applied to this triangle, is

$$\cos PM = \cos AP \cos AM + \sin AP \sin AM \cos PAM.$$

Substituting $PM = 90° - \beta$, $\alpha = AP$, and $h = AM$, and noting that $\angle PAM = \angle NAP - \angle NAM = 90° - (g - (k+\theta))$, yields

$$\sin\beta = \cos\alpha\cos h + \sin\alpha\sin h\sin(g-k-\theta). \tag{9.1}$$

Mayer says that, in principle, three observations suffice to solve for the three parameters α, β, θ, but that in practice this proves to be very difficult. Therefore he approximates the last equation as follows.

In triangle APM, we have $MP - PA < AM < MP + PA$, or equivalently $90 - \beta - \alpha < h < 90 - \beta + \alpha$. By inspecting his data set, Mayer could deduce from the difference of the minimum and maximum values of h therein, that $\alpha \approx 1°40'$.

31 'But to that end I have selected only those observations that are more correct than others and, because of the circumstance the moon was in at that time, more fit to the examination of the orientation of the lunar axis.' („*Ich habe aber dazu nur diejenigen Beobachtungen ausgelesen, welche vor andern richtig und wegen der Umstände, in welchen sich der Mond damals befunden hat, zur Untersuchung über die Lage der Mondaxe tauglicher sind, als die übrigen.*“) [Mayer, 1750a, p. 122].

32 These tables do not have to be of very high accuracy, since they here serve only to supply the *mean* longitude of the ascending node. The coefficients of its mean motion are computed easily and accurately from observations of eclipses one or more Saros periods apart. To find Manilius' longitude g, Mayer also needed the true ecliptic longitude of the moon, which he had to take from observations because the tables at his disposal were not accurate enough for that.

With this preliminary value for α, he could estimate $\beta \approx 14°42'$ and $\theta \approx 10°$ or so. Therefore he felt justified to take $\sin\alpha = \alpha$, $\cos\alpha = 1$, and $\sin\alpha\cos\theta = \alpha$. Expanding $\sin(g-k-\theta)$ and making approximations, the right hand side of (9.1) yields $\cos h + \alpha \sin h\,(\sin(g-k) - \cos(g-k)\sin\theta)$. On the other hand, introducing n such that $90° - \beta = h - n$, and consequently $|n| < |\alpha|$, he obtained $\sin\beta = \cos(h-n) \approx \cos h + n\sin h$. Together with the approximated right hand side, this gives, after division by $\sin h \neq 0$,

$$n = \beta - (90° - h) = \alpha\sin(g-k) - \alpha\sin\theta\cos(g-k). \tag{9.2}$$

This is a linear relation between the unknown parameters α, β, and $\sin\theta$, with numerical coefficients computed from the known angles g, h, and k. Values can still be expressed in degrees, the linearization of $\sin\alpha$ and $\sin n$ not withstanding, because the linearization introduces the same factor $\frac{180}{\pi}$ on both sides of the equation. The neglected second-order terms are of the order of 10^{-4}. Modern least-squares solution of the system of equations (9.1) with Mayer's data differs from the solution of the linearized system (9.2) by only a few seconds in α and β, and about $4'$ in θ. This is too small to affect Mayer's conclusions presented below; we will see that the estimated standard deviations of the three parameters are considerably larger.

9.5.2 The 'method' of averages

Now that we have this relation (9.2) between the observed and the unknown quantities, we will discuss Mayer's use of observations to find the unknowns. Mayer selected three observations with respectively a large, medium, and small value for h, to obtain three well separated equations in the unknowns α, β, θ:

$$\begin{aligned}
\beta - 13°\ 5' &= +0.9097\alpha - 0.4152\alpha\sin\theta,\\
\beta - 14°14' &= +0.1302\alpha + 0.9915\alpha\sin\theta,\\
\beta - 15°56' &= -0.8560\alpha + 0.5170\alpha\sin\theta.
\end{aligned}$$

Solving, he obtained $\alpha = 1°40'$, $\beta = 14°33'$, and $\theta = 3°36'$.

While in previous decades an astronomer would have been perfectly satisfied with this result and moved on to other business, Mayer remarked:

> But because errors are often to be supposed in the values of g and h that are deduced from observations, which [errors] are impossible to avoid, yet they have an influence on the values of α, β and θ: so also must we not completely trust the present determination, which is deduced from only three observations. One must only try three other observations to get convinced of this.[33]

33 *„Weil aber in den aus den Beobachtungen geschlossenen Größen von g und h manchmal auch Fehler zu vermuthen sind, die sich unmöglich vermeiden lassen, gleichwol aber in die Größen von α, β und θ einen Einfluß haben können: So dörfen wir auch der gegenwärtigen Bestimmung, die nur aus dreyen Beobachtungen hergeleitet worden, nicht völlig trauen. Man darf nur eine Probe mit dreyen andern Beobachtungen anstellen, um hievon überzeugt zu werden"* [Mayer, 1750a, p. 151].

To reduce the influence of the observational errors on the solution, Mayer's remedy is to take all his 27 observations simultaneously ('*zugleich*') into account, each yielding one of 27 near-linearized equations. To solve the resulting overdetermined system, he divided the equations in three classes of nine each. The first class held the nine equations with the largest positive values for $\cos(g-k)$, the second class those with the largest negative values for $\cos(g-k)$, and the remaining equations (i.e., those with large values for $\sin(g-k)$) went into the third class. Then he summed the nine equations in each class to obtain again three equations in three unknowns, with deliberately large differences between the coefficients:

$$
\begin{aligned}
9\beta - 118^\circ\ 8' &= +8.4987\alpha - 0.7932\alpha\sin\theta,\\
9\beta - 140^\circ 17' &= -6.1404\alpha + 1.7443\alpha\sin\theta,\\
9\beta - 127^\circ 32' &= +2.7977\alpha + 7.9649\alpha\sin\theta.
\end{aligned}
$$

He obtained the solution $\alpha = 1^\circ 30'$, $\beta = 14^\circ 33'$, and $\theta = -3^\circ 45'$.

The key idea of the method of averages is as follows. Divide the total corpus of equations in as many classes as there are unknowns, then sum the equations in each class. Solve the resulting system of equations. The solution is believed to be more accurate if more observations are employed, and if the division in classes is aimed at maximizing the differences between the coefficients of the final equations:

> But the advantage consists therein, that through the above division in three classes, the differences between the three sums become as large as possible. And the larger these differences are, the more correct are these unknown quantities α, β, θ to be found from them.[34]

In the case of this particular model, division in classes works surprisingly well to maximize the differences between the coefficients of the three summed equations. We will return to the reasons behind this later in the section. Mayer does not supply a reasoning why all classes should contain an equal number of equations.

Next, Mayer comes to a very interesting error estimation, based on the two different solutions just quoted. Comparing the two values found for α, he remarks that one is derived from nine times as many data as the other, which makes it 'nine times as good' and its (probable) error nine times less. He introduces x for the error in α, and writes $\alpha = 1^\circ 30' \pm x$; the first determination from only three observations yielded $\alpha = 1^\circ 40'$ so the error therein was then $10' \pm x$. Mayer expresses his supposition that the error behaves inversely proportional to the number of observations as $\pm x : \frac{1}{27} = (10 \pm x) : \frac{1}{3}$, for which he gives a solution[35] $x = 1'\frac{1}{4}$. The other solution, $x = 1'$, he does not mention; but he concludes that the true value of α might differ $1'$ or $2'$ from $1^\circ 30'$. Likewise, β must be about right and an error of 1° may exist in θ. Using a bootstrap technique[36] I established standard deviations for α, β, and θ

34 *„Der Vortheil aber bestehet darinn, daß durch die obige Absonderung in drey Classen die Unterschiede unter den dreyen Summen so groß geworden, als es möglich war. Denn je größer diese Unterschiede sind, je richtiger lassen sich die unbekannten Größen von α, β und θ daraus finden“* [Mayer, 1750a, p. 154].

35 The solution was erroneously printed as $1''\frac{1}{4}$ [Mayer, 1750a, p. 155].

36 [Press et.al., 1995, pp. 291–292].

of respectively $2.9'$, $2.6'$, and $2°7'$. Mayer recognized that the determination of θ is not very reliable because the angle formed by the lunar equator and ecliptic (i.e., α) is so small. I am not aware whether Mayer made any other error estimates of this kind.

Stigler[37] highlights the novelty of Mayer's data handling, and its influence on contemporary and later mathematicians (notably Lalande and Laplace). He stresses that Mayer's treatise is remarkable for its time, because Mayer found it useful to combine so many observations, and because he attempted a quantitative error estimate. Mayer was too optimistic, in our modern view, when he supposed that the error in his determination of α behaved inversely proportional to the number of observations; yet (as Farebrother pointed out) the exploitation of the fact that a relation between the two exists at all was an important step in the theory of errors.[38]

Stigler's investigation into the intellectual climate in which the method of least squares was conceived, leads him to a comparison of this work of Mayer's, with slightly earlier work of Euler's on the perturbations of Jupiter and Saturn, and with later work of Boscovich and Laplace. All these people were involved in the fitting of model parameters to observations, and they did so in more or less innovative ways. Stigler concludes that Euler, as a mathematician, was wary of the accumulation of *maximum* error when observations are combined. Mayer as a practising astronomer was aware (more than Euler) of the cancelling properties of random error. But, as Stigler stresses, Mayer did not go so far that he applied the property of cancellation in all cases; he allowed it only when similar data were taken under similar circumstances (i.e., same observer, same instrument, etc.). Stigler looked upon Mayer's division into three disjoint classes as a division among different observational circumstances, reflected in the coefficients, that Mayer preferred to keep separate. Later, Laplace would go further than Mayer by treating the set of observations as a whole. He devised a general method of combining observations, which, when applied to the Manilius data of Mayer, would combine the entire corpus of 27 equations by addition or subtraction in three different ways, to arrive at a system of three equations, each depending on all 27. In Laplace's method each of the three combined equations depended on all of the original equations instead of on a subset. From Euler, via Mayer, to Laplace, Stigler signals a steady increase of the willingness to let random observational errors cancel each other.

On the other hand, Stigler signals a lack of generality in Mayer's procedure. Mayer obtained good results because of his design of the experiment and because of the geometry of the problem (and, I would add, also because of his skill as an observer). One particular circumstance in Mayer's formulation of the libration problem was the appearance of both $\sin(g-k)$ and $\cos(g-k)$ as coefficients in the equations. Upon putting those with extreme values for the former in the first and second

37 I discuss Stigler's work [Stigler, 1986, Ch. 1]. Other commentaries on Mayer's *Umwälzung* tract are in [Forbes, 1980, pp. 48–52], [Wolf, 92, II, pp. 506–509], [Lalande, 1764, Vol. II, pp. 1234–43], [Farebrother, 1998, pp. 11–15], among other places.

38 [Farebrother, 1998, p. 15].

classes, those with extreme values for the latter are necessarily left for the third class. This is a particularly helpful relation between two of the three coefficients in the equation. It is not at all obvious how Mayer's procedure would generalize in the case of more unknowns or when the relations between the coefficients are less favourable. The special circumstances of the current application rendered the criterion for class allocation quite obvious. But how would he apply this procedure to the problem of adjusting two dozen lunar equAtions to over a hundred observations? As the number of unrelated coefficients grows, the allocation of the observations over just as many classes gets increasingly arbitrary. At the least it would be required to investigate the effect of the chosen allocation on the fitted parameters.

To understand why the division in classes is so effective in maximizing the differences between the coefficients of the three summed equations, we return to equation (9.2). It contains the three unknown quantities α, β, θ, which need to be determined. It also contains three known quantities g, h, k which depend on the time of observation. These three quantities yield the numerical coefficients of the equations, in the form of $90^\circ - h$, $\sin(g-k)$, and $\cos(g-k)$. It is immediately apparent that these three coefficients are not independent. In fact, Mayer's choice to fill classes I and II with those equations that have extreme values of $\sin(g-k)$, leaves equations with extreme values for $\cos(g-k)$ to constitute class III. These class III values all happen to have the same sign, for reasons that we will explore further down. If they had mixed signs, the equations in the third class would not sum to an equation with a large last coefficient. The situation is depicted in figure 9.2. But at the same time, figure 9.1 shows that the arc $90^\circ - h = MB$ is also governed by $g-k = NB$. It can be seen that MB is least when P is on the arc AMB and in between A and M; it is largest when P is on AMB but on the far side of A. Thus, the three numerical coefficients are not particularly independent.

Why do all class III equations have a positive coefficient in their last term? Stigler points out that this property is to an extent responsible for Mayer's success, but I disagree with Stigler when he suggests (pp. 22–23) that they turn out to be positive because of Mayer's choice of crater. In contrast, I will argue that they turn out positive because of the dates at which Mayer observed. In passing, however, we note that equations in class III with negative values for these terms could have been easily handled if Mayer were prepared to *subtract* them from the others in their class, instead of adding them.

Mayer's choice of the crater Manilius was governed by the necessity to select a distinctive feature close to the centre of the visible lunar disc, otherwise he would be unable to make accurate measurements of its position. Thus, from the vantage point of the moon's centre, the direction of the feature should not differ more than, say, 20° from the direction of the earth. As a consequence, the ecliptic longitude g of the feature is predominantly dependent on the ecliptic longitude of the earth, as seen from the moon; or conversely, on the longitude of the moon as seen from the earth. This quantity is periodic with period a (tropical) month.

The longitude of the ascending node changes approximately 20° during the full year of Mayer's Manilius observations, which makes it nearly constant with respect to g. We see that $g-k$ completes a circuit of the ecliptic in approximately one month.

Classes I and II have devoured those observations where $g-k$ is near 90° or 270°, leaving for class III the cases where $g-k$ is closer to 0° or 180°. When we realize that $g-k$ is approximately the arc ☊b in figure 9.1, it follows that class III holds those cases where the earth-moon direction is more or less along the axis ☊NCn☋, and the sign of $\cos(g-k)$ is positive or negative as an earth-bound observer views the moon in the direction of the descending or ascending node, respectively. We conclude that the coefficients $\cos(g-k)$ in the class III observations are negative (thus giving a positive last term in the equations, taking note of the minus sign in equation (9.2)) because Mayer observed them on dates when the moon was nearer its ascending node.

It would be illustrative to know whether Mayer planned these observing days in advance, or whether he selected a convenient set of observations *a posteriori*. Mayer mentioned ten extra observations of Manilius in his text, which he did not include in his working data set for various reasons: either they seemed to be less accurate, or they were inappropriate for the determination of the orientation of the lunar axis. I calculated the value of $g-k$ for those observations, whereafter it appeared that one or two might have ended up in class III with the wrong sign. Perhaps that was the reason why Mayer rejected them. The scatter of the values of $g-k$ (figure 9.2) suggests that Mayer might have put some planning in his observation schedule. However, a large proportion of the observations was made in July of 1748, during a lunation that ended in a solar eclipse, whereafter the next lunation incidentally offered a lunar eclipse. Mayer gave ample evidence that these eclipses had his full attention, in order to squeeze every possible bit of information from them.[39] That is why he was making more than casual observations of sun and moon around that time. And it just so happened that the solar eclipse occurred near the descending node, implying that most of his Manilius observations before and after the event were nearer the ascending node, when the crater was sunlit. Therefore, the values of $g-k$ populate predominantly the left half of figure 9.2, and Mayer had to wait half a year before he could make the rightmost observations in the figure, with the sun illuminating Manilius near the descending node. The unequal spread suggests seized opportunities rather than advance planning.

9.6 EULER'S LUNAR TABLES OF 1746

Did Mayer apply the method of averages to correct the coefficients of the lunar motion tables? It is unlikely that he did so after the advent of his spreadsheet tool

39 The same (and, unfortunately, only) volume of the *Kosmographische Nachrichten* contains an article where Mayer expounds his results [Mayer, 1750b]; see also the charts that Mayer and Lowitz drew in preparation of the events, probably working together, [Mayer, 1748] and [Lowitz, 1747].

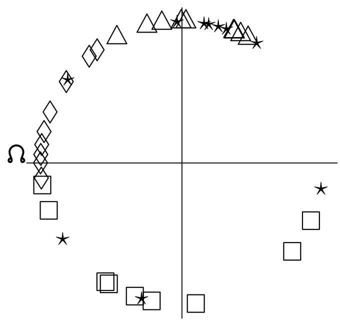

Figure 9.2: The values of $g-k$ plotted along the unit circle. The plot symbols are: $\triangle$ for Mayer's class I, $\square$ for class II, $\Diamond$ for class III, and $\star$ for the rejected data.

during 1753. The development of the spreadsheets begins in Cod. $\mu^{\sharp}_{41}$, a manuscript characterized in section 8.3 as extremely relevant to the various researches surrounding the publication of the `kil` tables. If Mayer had applied the method of averages *before* the spreadsheet tool, this manuscript would be the most likely place to find its traces. A search in Cod. $\mu^{\sharp}_{41}$ for systems of linear equations makes one halt at folios 273r–286v. Those folios form the object of study of the present section. These folios make up a complete quire, which I have already briefly alluded to on page 139.

The first few folios of the quire are clearly connected to Mayer's libration research: they contain calculations related to lunar craters (Manilius, Dionysius, and Menelaos) using observations taken in 1748, which Mayer included in the *Umwalzung* tract that we studied above. The following folios bear 37 lunar position calculations that employ the lunar tables of Euler, *Tabulae astronomicae solis et lunae*.[40] Mayer had prepared his own manuscript copy of Euler's tables; it is now in Cod. $\mu^{\sharp}_{14}$, fol. 1–13. Although Mayer copied the equAtions out of Euler's tables, he adjusted the mean motions from the Julian calendar and mean time on the Berlin meridian, to the Gregorian calendar and mean time on the Paris meridian. In some of his articles for *Kosmographische Nachrichten* he referred to these tables as the best available at the time.[41]

Clearly, these folios are out-of-sequence from a chronological point of view. Whereas the surrounding quires contain lunar position calculations based on the 1753 `kil` tables, this one was written perhaps four years earlier, when Euler's lunar tables were still the best available.

The 37 position calculations serve to compare the tables to observations of lunar meridian passages.[42] For each observation, Mayer deduces a linear equation in variables t, u, v, w, x, y, and z. For example for the observation of 12th February

40 [Euler, 1746].

41 See for instance [Mayer, 1750b, §12].

42 These observations are not related to the crater measurements above. Mayer copied most of the observations out of the *Mémoires* of the Paris Academy for the year 1739. In Cod. μ_{12} are his extracts for two-thirds of the data; I was unable to trace 12 observations of 1742 there.

1739 he derived:

$$1000t - 58u + 920z - 700v + 847w + 62x - 962y - 328000 = 0.$$

The constant term in all these equations is 1000 times the difference of computed and observed lunar longitude (here, $328''$). All equations have a term $1000t$, where t presumably represents a number of arc-seconds by which the epoch longitude of the moon should be adjusted. The variables t through z stand for changes to the coefficients of the tables. The coefficients of $v \dots z$ in the linear equations are the sines of the arguments of those tables, multiplied by 1000. t and u express adjustments to the mean motion of the moon's longitude and apogee, respectively.

So here we have Mayer deriving 37 equations in 7 unknowns, approximately in the same period when he worked out the orientation of the lunar axis with the method of averages. A solution (in an approximative sense) of the overdetermined system would provide a correction of Euler's lunar tables. Interestingly, in Mayer's manuscript copy of those tables, I found the following notes written at a later instance next to the table headings: mean motion table, $t = -40''$ and $u = -12'$; table I, $z = +6'40''$; table II, $v = +2'10''$; table III, $w = -57''$. That looks like a partial solution of the overdetermined system. Unfortunately I have not been able to locate the papers where Mayer calculated this partial solution. Numerical experiments on a computer showed that a least-squares solution is unstable and liable to drastic changes when one or more observations are discarded.

With so little evidence we can only speculate. We are not sure if Mayer corrected Euler's tables before or after he successfully applied the method of averages in his libration research, although the time interval between the two events is unlikely to be more than about a year. Around that time, Mayer remarked that he had made several adjustments to Euler's lunar tables, which supports this interpretation.[43] I hypothesize that Mayer endeavoured to apply the same method of averages, but that he immediately became aware of its limitations. With more than just a few variables, it is no longer a trivial matter to decide on the distribution of equations over classes. This limitation seems to be less prevalent in Laplace's generalization of the method of averages.

Mayer's interest in Euler's tables had disappeared before the end of 1750 (see page 27). Apparently, his attempts at a systematic improvement of the Euler tables had a longer lasting value to him. Mayer inspected his earlier work in connection with fresh position calculations made during 1753: after the publication of the `kil` tables, therefore even after the development of the spreadsheet technique. The dislocation of these older folios shows us that the method of averages was still on his mind when he had already invented the spreadsheet tool.

Interestingly, similar sets of linear equations as were presented above, in up to eight unknowns, figure in the final chapters of Euler's treatise on the great inequA-

43 '...but that I made some improvements [to Euler's tables] guided by many observations...' („...*daß ich aber aus Anleitung vieler Beobachtungen einige Verbesserungen gemacht habe...*“) [Mayer, 1750b, §12].

lity of Jupiter and Saturn.[44] There, Euler attempts to adjust his equAtions of the Saturn orbit to observations, and the linear equations play exactly the same role as in Mayer's work to improve on Euler's tables. Moreover, Euler and Mayer both apply themselves to the same kind of ad-hoc strategies to get an impression of the magnitudes of the unknowns. They concentrate on the unknowns with the largest coefficients first and neglect those with the smallest coefficients, altogether in a much more haphazard way than in Mayer's later method of averages. After having fixed some of the unknowns, Euler finally tabulates what happens to the differences between observed and computed positions of Saturn when he assumes values for one and then another one of the remaining unknowns. This makes one think of Mayer's spreadsheets, and it may indeed well be that he obtained the basic idea from this influential treatise of Euler.

44 [Euler, 1749a], pp. 123–141 of the reprint edition.