

ON THE FLEXURE OF MERIDIAN CIRCLES

R. d'E. Atkinson

(Communicated by the Astronomer Royal)

(Received 1955 March 25)

Summary

The methods which have been used or suggested for determining meridian-circle flexure are without exception unsatisfactory. Very few of them can detect a second harmonic, and many of the results actually obtained are unacceptable astronomically. The paper analyses in detail the effects of uncertainties in the flexure function, and also in the refraction, on the solar corrections, giving a more rigorous treatment than has so far been used; this rigour, and also an improvement in the methods of determining flexure observationally, both seem desirable if any further advance is to be made in the precision of the results. A new method of directly determining all harmonics of the flexure function is proposed, and the necessary equation of condition is set up for a particular case which seems the most suitable one.

Introduction.—It is universally recognized that the flexure of a meridian circle is one of the most difficult of instrumental errors to determine. As was first pointed out by Bessel (**1**), a value applicable to the horizontal position can be obtained immediately, by observing the two collimators after they have been set on one another; but the natural assumption that the flexure at other points on the circle will be this “horizontal flexure” multiplied by the sine of the zenith distance frequently proves irreconcilable with the astronomical results. Various methods of determining the true form of the flexure, as a function of zenith distance, have been suggested, and some have been tried very extensively, but none seems free from fairly serious objection; since any uncorrected error will enter directly into the declinations of every star catalogue based on the instrument in question, and will also affect the corrections to the Sun’s elements, the position is definitely unsatisfactory. The present paper reviews the different attacks which have been made on this problem, discusses in detail the effects on the solar elements which different correction functions will produce, and proposes a new method of measurement which appears to be free from the objections that have been explicitly noticed.

In order to avoid misunderstanding, it is first desirable to clarify a point of terminology. Flexure is sometimes spoken of as positive if it “increases the zenith distances”, whether these are north or south, and the increase is then meant in the numerical sense and not algebraically; it is, however, simpler in some ways to regard the flexure as an ordinary transcendental function of z , with z positive north of the zenith (say), and to agree that in all cases the observed Z.D. is to be increased *algebraically* by the amount of the flexure in order to get the true Z.D. Such flexure functions as $f = f_0 \sin z$ must then be called *anti-symmetrical*, not *symmetrical*, about the zenith, although they correspond to an

arithmetically (or even "geometrically") symmetrical effect; a flexure varying as $\cos z$ will similarly be spoken of as symmetrical about the zenith, although it deflects the zenith point, and although arithmetically it increases the Z.D.'s on one side of the zenith and decreases them on the other. In the present case, the algebraical approach seems definitely more convenient than the geometrical, and we shall follow this convention throughout.

The nature of "flexure".—It is well known that if the eye-end and objective-end of the telescope suffered identical flexures, their combined astronomical effect would be zero. The two tubes can be made nearly identical, especially if a draw-tube for focusing is dispensed with (2); if their lengths are exactly equal, and if the "length" (flange to second nodal point) of the lens cell is equal to the "length" (flange to focal plane) of the micrometer box, and if the instrument is also well balanced (as it normally is, with some delicacy) both with and without the eye-end and cell, interchanging the latter will eliminate the effect of any differential flexure of the bare tubes which may still exist, at least if this occurs only as a first harmonic (or as a sum of odd harmonics) in the circle readings; but the objective and micrometer themselves are of course quite dissimilar, and any internal movement in either will have the same astronomical effect as a flexure, with the added complication that it may possibly be non-elastic, and may even show discontinuities. For this reason, and also because it will not in any case eliminate any even harmonics, there seems little point in actually interchanging the two ends, and the practice has indeed been fairly generally abandoned, even in instruments which readily permit it. Personality errors should be eliminated by the systematic use of a reversing prism, but if this precaution is omitted they also will enter into the "flexure"; indeed, in at least one programme (3), a systematic difference in the zenith distances of near-zenith stars was found, depending on whether the observer lay head-north or head-south, even though a reversing prism was used, and it thus seems that this device cannot be relied on implicitly. Certainly there is no reason why such personality effects as are not eliminated should vary as $\sin z$, or should not have a discontinuity at the zenith.

The problem is (as it seems) further complicated by the possibility of flexure in the circle; the circle readings have to be combined with the Z.D. micrometer readings in most (but not all) flexure tests, and of course in all astronomical observations, and even if one assumes that circle flexure will vary as $\sin(z - \gamma)$ there may be no evident reason why γ should be small. If it is not, the combined flexure of telescope and circle will have terms in both $\sin z$ and $\cos z$; in the Washington 9-inch T.C., the "circle flexure" was found to be considerably larger than the other (3), and it seems that the resultant might have any phase.

It is however a little surprising that observable circle flexure should exist at all. The true flexure of an elastic but symmetrical circle, firmly attached to the main axis all round but free to droop under its own weight, would produce only the same effects as slight constant displacements of the various microscopes, and would thus be completely irrelevant even if only one microscope were used; the microscope displacements would not even be noticed, since the microscopes are initially adjusted on the basis of the circle divisions as actually read. If the circle is not equally stiff all round, or is not attached to the axis in a uniform and symmetrical way, detectable circle flexure may of course exist; but if local weaknesses are postulated, they seem in general more likely to affect the differences of the microscope-readings than the sum of all four (or six), and on the whole

one would then be inclined to expect observable circle-flexure to occur as a fairly high harmonic of z , not merely as the first, which is what has been found in practice. A single local excess of weight would be effective, but it seems somewhat improbable that any manufacturing flaw of this sort could be large enough. There is, however, one case where an apparent first harmonic would be found alone: a uniform circle which was effectively attached to the axis at only two points, with one of those attachments stiffer than the other, would suffer rotations about some virtual pivot (the resultant of these two partial constraints), and if the constraints were elastic these rotations would be proportional to the sine of that pivot's Z.D., and would be equivalent to translations plus similar rotations about the axis; the translations would disappear with two or more microscopes, and the rotations would be interpreted as first-order flexure. The importance of this lies in the fact that the normal method of determining circle flexure is to have the circle "movable" (i.e. with respect to the telescope), to rotate it every other year (say) to a new position, and to analyse the measurements of horizontal (and perhaps also vertical) flexure in terms of these circle positions; it may, however, be much harder to guarantee a fully adequate and symmetrical connection between circle and axis when this provision for rotation is made than when the circle is deliberately constructed as a "fixed" one. It would thus not be altogether surprising if first-order circle flexure were an error introduced mainly by the provision made for detecting and measuring it, and a well-designed "fixed" circle might in that case be preferable to any attainable excellence in a movable one.

Methods for determining flexure.—(a) *Bessel's method.* Since the combined flexure of circle and telescope is what is required astronomically, the ideal would be to determine this directly at a number of different zenith distances. In principle, as again Bessel first pointed out, it would be possible to set up a pair of collimators in some other position than the horizontal, or in a series of such positions, to determine the flexure for each position, and to express the results as a function of z . This method could not be employed, as it stands, in an ordinary pavilion, and does not appear to have been explored at all; but it would seem reasonably practicable if the lower of the two collimators were replaced by a plane mirror and the upper one were fitted with a Bohnenberger eyepiece. However, even if suitable arrangements could be made, the collimator method does not really give the true astronomical flexure in any case, but only the difference of two values 180° apart; thus however many collimator positions one uses it cannot detect any even harmonics in the flexure function, and in particular it cannot detect a second harmonic. (This weakness seems not to have been pointed out previously.) If it can be assumed that there are no harmonics above the first, there is no real need for a large number of collimator positions after all; a single pair of vertical collimators, in addition to the horizontal pair, is sufficient to determine the complete function, and the lower one of the vertical pair can then be replaced by the ordinary mercury-pool (4).

The "vertical flexure", i.e. the flexure (including that of the circle) evaluated by the use of vertical collimators or their equivalent, is not really material in a modern instrument, since it is eliminated by reversal. Indeed, if the flexure function is expressed as a Fourier series in z , this is true of all the cosine terms in the complete series. If there are no sine harmonics above the first, it must thus be sufficient to determine the horizontal flexure f_0 by means of the collimators, reverse the instrument regularly, and assume that $f = f_0 \sin z$; rotating the circle

in addition should be unnecessary, so far as this question goes, whether there is circle flexure or not. If the circle is never moved, many of the corrections discussed below can also be obtained with greater weight; and since the only information which is obtained by moving it is information that is not wanted, it would seem better to keep it in one position even if it is designed to be movable.

(b) *Loewy's method.* Methods of test of this kind, i.e. employing only collimators, mirrors, etc., but not stars, may be classed as "purely instrumental"; there may also be "purely astronomical" methods (or, at least, checks); and there are methods which we may call mixed "instrumental-astronomical" ones. Two other purely instrumental methods seem worthy of mention here. The first is due to Loewy (5); a meniscus lens, with its concave radius of curvature equal to half the length of the telescope and its focal length equal to half that radius, was very firmly mounted in the central cube, with the concave side towards the eyepiece. Using the concave side as a mirror, a real image of the Z.D. micrometer wire could be seen in the same focal plane, and the setting for coincidence noted; the variation of this setting, as the telescope was pointed to various zenith distances, gave the flexure of the eyepiece half of the telescope. Using the lens property, a mark on the telescope objective could also be imaged in the focal plane; and the variation of Z.D. micrometer settings on this, as the telescope was moved, gave the sum of the flexures of the two halves. The difference of the two then followed. The main difficulty, that of determining the flexure of the auxiliary lens and mounting itself, was solved (it was claimed) by adding weights to this mounting, repeating the observations, and "eliminating" the flexure due to the original weight. If the difficulty certainly can be got over in this way, the method appears sound; but in the case of the reflection observations it is tilts of the mounting, even more than bodily displacements, which are to be feared, and it might be difficult to be sure that the loads were always quite suitably applied, from this point of view. Even a small miscalculation here can have very serious effects indeed. The results actually obtained were claimed to be very consistent, and it is interesting to note that they did indicate the presence of an appreciable second harmonic; the flexure was found to be

$$F = +1''.04 \sin z + 0''.17 \cos z + 0''.30 \sin 2z - 0''.23 \cos 2z. \quad (1)$$

The method evidently yields only the telescope flexure, without the circle flexure. However, it was argued that the latter would be expected to be zero; and the horizontal collimators, which, of course, measure the combined (horizontal) flexure, gave $+1''.15$, in fairly good agreement with the $+1''.04$, above, for the telescope only. This is perhaps the only case in the literature where an important second harmonic was directly determined; but the possibility that some flexure of the mounting was not fully accounted for by the process of varying the loading cannot be entirely ruled out. The method is inconvenient, since the meniscus prevents all astronomical observations while it is in place, and it does not seem to have been adopted elsewhere.

(c) *Bonsdorff's method.* The other purely instrumental method is due to Bonsdorff (6); a plane mirror was supported in front of the objective, by a mounting designed to guarantee that flexure of the mounting would move the mirror strictly parallel to itself, while flexure of the telescope would not affect it at all. Auto-collimation measurements were then made with this mirror, in a variety of telescope positions, so as to derive the complete flexure function of the

telescope. It was stated that adding extra weights to the mirror verified that it was functioning correctly, the reading at any one Z.D. being unaffected, as it should. It was also stated that the results indicated a variation purely with $\sin z$, but this appears to be incorrect; the treatment given to the individual readings would necessarily have removed any term in $\cos z$, and if these readings themselves are examined they indicate not only that there was such a term, but that it was actually about three times as large as that in $\sin z$. It is unfortunate that no comment was made on this; it seems rather improbable that so large a $\cos z$ term could be real, for in this case also the circle flexure evidently does not enter. The method is believed to have been tried at several other observatories, but to have been found difficult in most cases. It is clear that perfection of design, and high precision in manufacture, are essential if the mirror's movements are to be sufficiently parallel; a slight difference in stiffness as between two members that were nominally identical, or a slight difference in friction as between two joints that were both nominally free, would be enough to impair the action, and it would perhaps not be very surprising if two superficially indistinguishable mountings gave somewhat different results.

With both these methods, if the results did not agree with those obtained from the collimators, it would be difficult to know what to do with them. The difference might indeed be ascribed to circle flexure; but since the circle readings must be used in the astronomical work itself this would mean that the collimator results had to be relied on after all, and the others rejected (so far, at least, as the first harmonic is concerned), unless the form of the circle flexure could be discovered independently.

(d) *The "R. and D. method"*. By "mixed" instrumental-astronomical methods we understand those in which stars are observed, but nothing is assumed about their positions except the fact that for any one star the (mean) declination is constant; its absolute value does not enter. In addition, an instrumental arrangement is made which permits the same star to be observed in two or more different positions of the telescope, related to one another by some condition which is known with precision; and a comparison of the circle readings actually obtained, with those to be expected on the basis of the known condition, gives a condition which the flexure must satisfy. The best-known of these methods is that of observing the star both directly and by reflection in a mercury pool. The precisely-known condition in this case is that the sum of the two observed circle readings, together with the appropriate flexures, must be exactly 180° ; moreover, half the difference of the two readings should give the star's altitude, without relying on the ordinary nadir pool at all. Since the nadir observation is of a different kind from all observations of stars (and is also made in a different position of the observer, and of the micrometer), this possibility of eliminating any systematic error that may affect it appears at first sight valuable.

The method was tried extensively at Greenwich (7), but the results were never very satisfactory. It seems probable that one reason for this lay in the unsuitable nature of the pavilion in this particular case (8). (The aperture of the Airy T.C. pavilion is only 3 ft. wide, and is 16 ft. above the axis of the instrument; and the instrument itself is also partly sunk in a pit, so that the mercury pool is in many cases below ground level, which may perhaps add to the refraction dangers inherent in the pavilion.) The method appears to have been found more successful at Washington (9), though a mercury pool can of course be used with a wide

pavilion-aperture on very still nights only. It has not been employed recently at all; and actually, even in ideal circumstances, it is defective in principle: it cannot detect or eliminate any component of flexure which is antisymmetrical about the horizon. This weakness does not appear to have been pointed out in the literature, at least in this precise form; it seems to deserve explicit consideration.

Let c be the circle reading, reckoned from the zenith, positive for Z.D. north, when the telescope is set on a star at zenith distance z (as refracted). (For convenience, we speak of the combination of Z.D. micrometer reading and circle reading as the "circle reading c ".) Let $f(c)$ be the flexure at the circle reading c , positive if its effect is such that the true Z.D. is greater (algebraically) than c by the amount of the flexure; so that the zenith distance of a star (including refraction) is

$$z = c + f(c). \quad (2)$$

Let c_D be the reading for a given star if it is observed directly, and c_R the reading if it is observed by reflection in the mercury pool; c_R is supposed corrected for the necessary small change of latitude of the pool, though this precaution has perhaps not invariably been observed in practice.

Then we have

$$z = c_D + f(c_D) \quad (3)$$

and

$$z = \pi - c_R - f(c_R). \quad (4)$$

Let f now be divided into a portion, f_S , symmetrical about the horizon, and a portion, f_A , antisymmetrical about it; then $f_S(c_R) = f_S(c_D)$ and $f_A(c_R) = -f_A(c_D)$. Inserting these in (3) and (4), and adding and subtracting the results,

$$2z = c_D - c_R + \pi + 2f_A(c_D) \quad (5)$$

and

$$0 = c_D + c_R - \pi + 2f_S(c_D). \quad (6)$$

Equation (5) shows that any systematic error affecting the zenith (or the nadir) is indeed eliminated when the true Z.D. is inferred in this way (since it will enter equally into c_D and c_R); and equation (6) is an equation of condition for determining f_S ; we may assume for example

$$f_S(c) = \Sigma [a_{2n} \cos 2nc + b_{2n+1} \sin (2n+1)c] \quad (7)$$

using as many terms as seem reasonable, and on taking a sufficient number of observations we shall be able to solve by least squares for the a 's and b 's and so determine f_S completely. But there is no similar equation of condition for f_A , and f_A is present in (5) and also, of course, in (3) and (4): thus one cannot say that "the combination of a direct observation with a reflected one eliminates the effect of flexure", nor can one infer a total flexure for use with single direct observations (which must be the majority), unless f_A is known on other grounds.

The terms which are antisymmetrical about the horizon are, of course, the odd cosines and the even sines of c (or, as we may equally well say, of z); neither odd nor even sines are cancelled out by reversing the instrument. Thus even if the instrument is reversible the mercury-pool method will leave undetected and operative any term in $\sin 2z$ (or in $\sin 4z$, etc.) which may exist; and in addition it will fail to detect or eliminate any term in $\cos z$ (or in $\cos 3z$, etc.) if the instrument is not reversible, as the Airy T.C. itself was not. Higher harmonics than the second are perhaps unimportant, at least unless they occur in the circle rather than in the telescope, or are used to describe discontinuities, but it seems a little dangerous to assume that even the second harmonics are absent. It is an unfortunate coincidence that neither the collimator method nor the mercury-pool method can

say anything about a term in $\sin 2z$, nor can such a term be eliminated by interchanging the lens and eye-end, or by reversing the instrument.

(e) *The "astronomical method" (flexure and refraction).* In practice, no matter what methods are used for determining flexure to begin with, they are fairly often discarded again at a later stage of the same programme, the final choice of a flexure function being then based exclusively on the astronomical criterion that the distance from the pole to the equator is 90° . This is, of course, sufficient to determine one parameter only, and the flexure is therefore assumed to be of the form $f = f_0 \sin z$, with f_0 so adjusted that the astronomical criterion is satisfied. All preliminary determinations of flexure, e.g., with the collimators, are thus reduced to the status of mere devices for keeping the arithmetic compact; an accurate determination is pointless, except perhaps as a check that no unexplained change has occurred, and any arbitrarily assumed value of f_0 near the final one would be sufficient. Occasionally, when the original determinations of horizontal flexure and of refraction led to positions of the pole and equator which, together, did not satisfy the astronomical criterion, the flexure determinations have been kept and the refraction ones amended instead; but in either case the situation so created is rather unsatisfactory. It is always undesirable to throw away apparently good observations without discovering why they are wrong, and particularly so if it is only done in order to preserve others which may themselves be erroneous. The Washington 6-inch results for 1925-41 (10) and 1941-9 (11), which can certainly be classed as among the best available, show this difficulty very plainly; at the latitude of Washington, flexure and refraction are definitely hard to separate, on the basis of above-pole and below-pole declination comparisons, and when the distance from pole to equator proved to need correction it was necessary to sacrifice either the preliminary determinations of flexure or those of refraction with little to go on. In the former series, the refraction coefficient was changed by about $0''.16$ although the probable errors of the rejected values (slightly different at different dates) had been $\pm 0''.016$; in the latter series, the flexure coefficient was changed by $0''.04$ while the probable errors of the rejected determinations (12) were $\pm 0''.03$. The conflict in the second case is superficially less, but only because the equator-point error which had to be corrected was much less; if it had been as large as the other, the flexure coefficient would have had to be changed by $0''.23$.

In both cases, of course, everything that was known about possible systematic uncertainties, as opposed to the visible probable errors, was considered before making the decision; but it does now appear at least formally possible that the provisional determination of refraction and the provisional determination of first-order flexure may both have been substantially correct, and that the error of the equator-point may have been due to a second harmonic in the flexure, which could not have been detected by the methods used. And at least it is clear that a considerable advance in the reliability of the final results might be hoped for, if the different possibilities could be more surely separated. It is the case, unfortunately, that even though two alternative correction functions may not be clearly distinguishable over a moderate range of near-polar stars, they may have appreciably different runs over the 47 degrees of declination, centred on the equator, which are involved in solar observations; planetary observations necessarily involve a somewhat larger range still, and the range of declinations observed in preparing a star catalogue may be larger again. It seems worth while now to investigate just how serious these effects may be; we shall see that they can in fact appreciably exceed the probable errors claimed for some of the solar corrections.

Analysis of the solar corrections.—It is well known that if the residuals (Obs.-Tab.) of the Sun's declination are analysed, they yield corrections to the Sun's mean longitude L , to the obliquity ϵ , and to the two quantities h and k from which the eccentricity and the longitude of perigee can be inferred; in addition, there usually remains a constant term, $-\Delta\delta_0$ (13), which is commonly regarded as "the error of the equator-point". $\Delta\delta_0$ would be zero if the instrumental errors, and the refraction, were correctly known and allowed for; if it is not zero, and is to be removed, this must be done by applying some specifically assumed correction-function, such as

$$\Delta_c\delta = p + q \sin z \quad (8)$$

(if the first harmonic of the flexure is held to be responsible) or

$$\Delta_c\delta = p + q \tan z \quad (9)$$

(if the refraction is); as we have now seen, it may be desirable to consider also

$$\Delta_c\delta = p + q \sin 2z. \quad (10)$$

In each case, if the latitude has already been so adjusted that declinations above and below pole are consistent, the constants p and q are to be so chosen that the function vanishes at the pole, i.e. we must have

$$\Delta_c\delta = K(\cos \phi - \sin z) \quad (11)$$

$$\text{or} \quad \Delta_c\delta = K(\cot \phi - \tan z) \quad (12)$$

$$\text{or} \quad \Delta_c\delta = K(\sin 2\phi - \sin 2z), \quad (13)$$

where ϕ is the latitude; and in current practice K is then determined so as to make

$$\Delta_c\delta(E) = \Delta\delta_0 \quad (14)$$

if $\Delta_c\delta(E)$ is the "equatorial" value of $\Delta_c\delta$, i.e. its value when $\delta = 0$. Whichever of the functions (11), (12), or (13) is selected, it will act as a systematic correction to the $D\delta_0$ which characterize a star catalogue obtained with this instrument (as compared with others), and this has indeed been the principal purpose of deriving it; in addition, as was remarked by Hough (14), it will in general affect the solar corrections too.

The equation of condition in its usual form is

$$-\Delta\delta_0 + \sin \epsilon \cos \alpha \Delta L + \sin \alpha \Delta \epsilon + 2 \cos \alpha \sin \delta \Delta h - 2 \sin \epsilon \cos^2 \alpha \cos \delta \Delta k = \Delta\delta \quad (15)$$

where $\Delta\delta$ are the declination residuals; and since the coefficient of Δk never changes sign it is clear that $-\Delta\delta_0$ is not strictly the mean value of these residuals; it is also not the value of the left-hand side of (15) when $\delta = \alpha = 0$; and neither of these values is strictly the one to which (11), (12), or (13) ought to be adjusted. Although, as we shall see, the practical error is not as yet really serious, it might easily become so if experimental accuracy improved a little further; we will therefore now establish what the rigorous condition for $\Delta_c\delta$ is.

We may eliminate δ from (15) by means of the relation

$$\tan \delta = \tan \epsilon \sin \alpha; \quad (16)$$

if we also insert the numerical value of ϵ , we obtain

$$\begin{aligned} & -\Delta\delta_0 + 0.398 \cos \alpha \Delta L + \sin \alpha \Delta \epsilon + (0.413 \sin 2\alpha - 0.010 \sin 4\alpha) \Delta h \\ & - (0.389 + 0.398 \cos 2\alpha + 0.009 \cos 4\alpha) \Delta k = \Delta\delta \end{aligned} \quad (17)$$

(neglecting higher harmonics); and this form shows explicitly that the coefficients are for the most part very conveniently orthogonal if we can assume equal weights

at equally-spaced values of α (15). We shall make this assumption for the present; the question of weighting will be further considered at a later stage. An additional simplification, so far as $\Delta_c \delta$ is concerned, follows from the fact that any function depending only on z is the same in the autumn as in the spring, at the same values of δ ; since (for the Sun) α and δ are connected by (16), we may express $\Delta_c \delta$ by a Fourier series in α , and it must therefore be by one that is symmetrical about $\alpha = \pi/2$ and $-\pi/2$. It can thus only involve the even cosines and odd sines of α , and these are all orthogonal to the coefficients of both ΔL and Δh , so that it is only the quantities $\Delta \delta_0$, $\Delta \epsilon$, and Δk which can be affected by the kind of correction we have to consider.

If we difference two least-squares solutions, both having identical coefficients on the left, and employing "observational" quantities on the right which differ by the correction function $\Delta_c \delta$, the resulting equations are exactly what we would get if we solved the equation

$$-\Delta_c \delta_0 + \sin \alpha \Delta_c \epsilon - (0.389 + 0.398 \cos 2\alpha + 0.009 \cos 4\alpha) \Delta_c k = \Delta_c \delta \quad (18)$$

as if it were an equation of condition, to obtain the three additional corrections $\Delta_c \delta_0$, $\Delta_c \epsilon$, and $\Delta_c k$ by the regular least-squares process. In the "corrected" solution, these will be added to the original $\Delta \delta_0$, $\Delta \epsilon$, and Δk , which have already been obtained by solving the "uncorrected" equation (17); and the rigorous criterion for the choice of the magnitude (but not, of course, the form) of the correction function $\Delta_c \delta$ is that there shall then be no absolute term left, i.e.

$$\Delta_c \delta_0 + \Delta \delta_0 = 0. \quad (19)$$

Writing $\Delta_c \delta$ as a Fourier series in α (with the restriction already noted), i.e. putting

$$\Delta_c \delta \equiv c_0 + c_2 \cos 2\alpha + c_4 \cos 4\alpha + \dots + d_1 \sin \alpha + d_3 \sin 3\alpha + \dots \quad (20)$$

in (18), we may readily derive the normal equations

$$\Delta_c \delta_0 + 0.389 \Delta_c k = -c_0 \quad (21)$$

$$\frac{1}{2} \Delta_c \epsilon = \frac{1}{2} d_1 \quad (22)$$

$$0.389 \Delta_c \delta_0 + (0.389^2 + 0.0792) \Delta_c k = -0.389 c_0 - 0.199 c_2 - 0.0045 c_4; \quad (23)$$

and separating (21) and (23)

$$\Delta_c k = -2.51 c_2 - 0.057 c_4 \quad (24)$$

$$\Delta_c \delta_0 = -c_0 + 0.977 c_2 + 0.022 c_4. \quad (25)$$

Equation (19) thus leads to

$$c_0 - 0.977 c_2 - 0.022 c_4 = \Delta \delta_0 \quad (26)$$

as the rigorous condition for adjusting the "amplitude" of the selected $\Delta_c \delta$.

The "equatorial" value of $\Delta_c \delta$ is given by

$$\Delta_c \delta(E) = c_0 + c_2 + c_4, \quad (27)$$

and it is, of course, $-\Delta_c \delta(E)$ which is strictly "the error of the equator-point". The traditional criterion for adjusting $\Delta_c \delta$, namely (14), is clearly not identical with (26).

We may replace (26) by an approximation which eliminates the need for explicitly deriving the Fourier series (20). The summer-solstice value of $\Delta_c \delta$, $\Delta_c \delta(S)$ say, is

$$\Delta_c \delta(S) = c_0 - c_2 + c_4 + d_1 - d_3 \quad (28)$$

while the winter-solstice value is

$$\Delta_e \delta(W) = c_0 - c_2 + c_4 - d_1 + d_3; \quad (29)$$

we thus have

$$\frac{1}{2}[\Delta_e \delta(S) + \Delta_e \delta(W)] = c_0 - c_2 + c_4, \quad (30)$$

and the right-hand side of this is very similar to that of (26); indeed, if c_2 and c_4 happen to be in the same ratio as the coefficients of $\cos 2\alpha$ and $\cos 4\alpha$ in (17), i.e. if $c_4 = +0.023c_2$, (30) becomes

$$\frac{1}{2}[\Delta_e \delta(S) + \Delta_e \delta(W)] = c_0 - 0.977c_2$$

while (26) becomes

$$\Delta \delta_0 = c_0 - 0.977c_2,$$

so that the correct adjustment of $\Delta_e \delta$ is obtained when

$$\frac{1}{2}[\Delta_e \delta(S) + \Delta_e \delta(W)] = \Delta \delta_0. \quad (31)$$

This equation may be regarded as a "rule of thumb" condition for adjusting $\Delta_e \delta$, and it will always be a fairly good one, since c_4 and 2.3 per cent of c_2 will both be small even if they are not equal; thus, so long as we rely on declination measurements only, we may say that the correct adjustment of $\Delta_e \delta$ is obtained by making the mean of its two "tropical" values, and not its equatorial value, equal to $\Delta \delta_0$. To take a particular case, if we wish to remove a $\Delta \delta_0$ of $+0''.40$ (say) by means of a $\Delta_e \delta$ of the form (12), in latitude 40° , use of the equatorial value will lead to

$$\Delta_e \delta = 0''.235 - 0''.197 \tan z \quad (32)$$

while the tropical mean leads to

$$\Delta_e \delta = 0''.204 - 0''.171 \tan z; \quad (33)$$

the difference amounts to $0''.03$ at the zenith, to $0''.05$ at the equator, and to $0''.10$ at $\delta = -30^\circ$, and the two resulting star catalogues would differ systematically by a $D\delta$ of this magnitude. Actually, as we shall see below, these differences should normally be reduced by anything up to 50 per cent; but the traditional use of (14) still appears a little undesirable, more especially since the amount of labour which it saves is trivial.

It may be useful to tabulate the results of applying the harmonic analysis (20) to the three functions (11), (12), and (13), for three representative ϕ -values. The five coefficients of (20), all normalized to $c_0 - 0.977c_2 - 0.022c_4 = +1$, i.e. adjusted to cancel a $\Delta \delta_0 = +1''.00$, and the corresponding values of K to insert in (11), (12), and (13), are shown in Table I, and values for other latitudes may be interpolated graphically.

TABLE I

	Correction (11) (sin z)			Correction (12) (tan z)			Correction (13) (sin $2z$)		
	$\phi=30^\circ$	40°	50°	30°	40°	50°	30°	40°	50°
c_0	1.015	1.019	1.023	0.969	0.933	0.868	1.092	1.092	1.092
c_2	+ .016	+ .020	+ .023	- .032	- .066	- .135	+ .094	+ .094	+ .094
c_4	+ .001	+ .001	+ .001	+ .000	+ .003	+ .011	+ .004	+ .004	+ .004
d_1	- .266	- .231	- .194	- .247	- .352	- .484	- .262	- .080	+ .080
d_3	- .006	- .005	- .004	+ .004	+ .013	+ .038	- .011	- .003	+ .003
K	+ .755	+ .738	+ .743	+ .406	+ .428	+ .366	+ .687	+ .604	+ .604

We may use these figures to study the effect of removing any particular $\Delta\delta_0$ by any one of the three assumed correction functions. If, for example, observations in latitude 50° have yielded a $\Delta\delta_0$ of $+0''.30$ (after correcting the pole), and it is decided that an erroneous first harmonic in the flexure is responsible, then since $0''.30K = 0''.223$ in this case, the correction function is

$$\Delta_e\delta = +0''.223(\cos\phi - \sin z) \quad (34)$$

and we see from (22) and (24) that the solar corrections previously obtained will be modified by the subsidiary corrections

$$\Delta_e\epsilon = 0''.30d_1 = -0''.06 \quad (35)$$

$$\text{and} \quad \Delta_e k = -0''.30(2.51c_2 + 0.06c_4) = -0''.02. \quad (36)$$

If we decided to correct the same $\Delta\delta_0$ by means of a second harmonic in the flexure, we should have similarly

$$\Delta_e\delta = +0''.181(\sin 2\phi - \sin 2z) \quad (37)$$

$$\Delta_e\epsilon = +0''.02 \quad (38)$$

$$\Delta_e k = -0''.07, \quad (39)$$

while the decision to correct it by modifying the refraction instead (16) would result in

$$\Delta_e\delta = +0''.110(\cot\phi - \tan z) \quad (40)$$

$$\Delta_e\epsilon = -0''.15 \quad (41)$$

$$\Delta_e k = +0''.10. \quad (42)$$

The differences between these solutions appear appreciable, in view of the fact that the internal probable errors claimed for $\Delta\epsilon$ and Δk may be of the order $\pm 0''.03$ or less, and it is somewhat embarrassing if one has to choose between them arbitrarily.

Table I also allows us to put the effects of an inability to decide between different correction functions in another way. For any given latitude, we may difference the figures for two different correction formulae (or may make any linear combination of all three such that the sum of the three amplitudes is zero), and the result will be an expression which will leave the separation between pole and equator unchanged. For lat. 40° , for example, subtracting the "sin z " figures from the "tan z " ones and multiplying by $\Delta\delta_0 = +1''.000$, we get

$$\begin{aligned} \Delta_e\delta &= +0''.428(\cot\phi - \tan z) - 0''.738(\cos\phi - \sin z) \\ &= -0''.055 - 0''.428 \tan z + 0''.738 \sin z, \end{aligned} \quad (43)$$

and this expression can of course be multiplied by any other amplitude-constant instead without altering its property of leaving the pole-equator separation unchanged. (It vanishes at the pole, of course, but it does not and should not vanish at the equator; for this reason it might be a little difficult to construct it without using the above harmonic analysis.) By direct trial, one can see how large an amplitude it may be given without seriously affecting the below-pole minus above-pole declination differences, $\delta' - \delta$, on which the adjustment of the pole itself depended; if the original choice, as between a flexure adjustment and a refraction adjustment, was demonstrably wrong, the application of (43), with a suitable (perhaps negative) amplitude, will result in an improved run of these $\delta' - \delta$, but even when this does not happen, the run will not be made worse over an appreciable range of amplitudes. This is, of course, merely a quantitative

version of the qualitative statement that flexure and refraction are not clearly separable. (In general, applying (43) will also slightly change the pole itself, owing to the fact that a finite range of circumpolar declinations is involved; this effect can be dealt with by a further approximation if necessary.)

The harmonic analysis of (43), taken directly from the appropriate coefficient-differences in Table I, is

$$\Delta_c \delta = -0''.086 - 0''.086 \cos 2\alpha + 0''.002 \cos 4\alpha - 0''.121 \sin \alpha + 0''.018 \sin 3\alpha, \quad (44)$$

and the corresponding solar corrections, from (22), (24), and (25), are (for this amplitude, i.e. $1''.00$)

$$\Delta_c \epsilon = -0''.12, \quad \Delta_c k = +0''.22, \quad \text{and (necessarily)} \quad \Delta_c \delta_0 = 0;$$

even though the largest amplitude likely to be permissible (so far as the circumpolars go) is probably not over $0''.3$ or $0''.4$, the uncertainty thus introduced is somewhat unwelcome.

Actually, the above analysis is an over-simplification. In general, corrections to the solar elements are derived from residuals in α as well as from those in δ , and it is usual to form combined normal equations and treat both sets of residuals together. The equation of condition for the right ascensions, if (as is usually done) we give equal weights to all $\Delta\alpha$ rather than to all $\Delta\alpha \cos \delta$, is

$$\begin{aligned} -E + \cos \epsilon \sec^2 \delta \Delta L - \cos \alpha \tan \delta \Delta \epsilon + 2 \sin \alpha \sec \delta \Delta h \\ - 2 \cos \alpha \sec \delta \cos \epsilon \Delta k = \Delta \alpha, \end{aligned} \quad (45)$$

and if we eliminate δ as before this becomes

$$\begin{aligned} -E + (1.004 - 0.086 \cos 2\alpha) \Delta L - 0.217 \sin 2\alpha \Delta \epsilon \\ + (2.136 \sin \alpha - 0.044 \sin 3\alpha + \dots) \Delta h \\ - (1.877 \cos \alpha + 0.042 \cos 3\alpha + \dots) \Delta k = \Delta \alpha. \end{aligned} \quad (46)$$

This equation shows much the same kind of orthogonality as (17); and when they are combined, with equal weights for equal intervals of α as before, almost all the non-diagonal terms in the normal equations still vanish. If we add (17) and (18) and solve them together with (46) in this way, remembering that (17) and (46) themselves have already been solved together, and giving equal weights to an observation in R.A. and to one in Dec. (as is usual), we find that equations (21), (22) and (23) reappear with their right-hand sides and the terms in $\Delta_c \delta_0$ unchanged; the coefficient of $\Delta_c \epsilon$ in (22) is increased from 0.500 to 0.524, and the coefficient of $\Delta_c k$ in (23) is increased from 0.231 to 1.994. Equations (25), (22) and (24) are thus replaced by

$$-\Delta_c \delta_0 = c_0 - 0.042c_2 - 0.001c_4 = +\Delta \delta_0 \quad (47)$$

$$\Delta_c \epsilon = 0.954 d_1 \quad (48)$$

$$\Delta_c k = -0.108c_2 - 0.002c_4. \quad (49)$$

Table I can readily be re-normalized by dividing each column by the value of $c_0 - 0.042c_2 - 0.001c_4$ appropriate to that column. The " c_0 " which result are now all very close to 1.0, while the new " c_2 " (and " c_4 ") are not substantially altered; thus the rigorous criterion (47), for adjusting $\Delta_c \delta$, which replaces (26), is now only slightly different from

$$c_0 = \Delta \delta_0, \quad (50)$$

and falls almost half-way between (14) and (31). $\Delta_c k$ is in all cases now very small, while $\Delta_c \epsilon$ is changed less than 15 per cent from its previous value; these last two statements reflect the fact, well known qualitatively, that in a "combined"

solution Δk is determined mainly by the α -residuals and $\Delta\epsilon$ by those in δ . Except where the original c_0 (Table I) differ markedly from unity, such difference-equations as (43) and (44) reappear without any major modification, and the star corrections are also little affected.

The above conclusions have been obtained on the basis of equal weights for normal points, both in α and δ , which are evaluated for equally-spaced α -values. Weights are sometimes assigned in proportion to the numbers of observations made in these ranges, and this will in general involve reduced weights for the winter months, and will somewhat upset the orthogonality and separability which have much facilitated the present analysis. For any moderately uniform distribution of observations, the results will not be very greatly altered; but there are additional weighting considerations, some of them difficult, and the total amount of uncertainty which these import may be appreciable.

In the first place, there is some evidence that residuals in α (seconds of arc) should be replaced by residuals in $\alpha \cos \delta$; this would reduce the weights near both solstices systematically, but not by very large amounts. Secondly, the usual assumption that observations in α and δ have equal weights is probably based on little more than the difficulty of coming to any decision at all on really objective grounds. The systematic errors of the two methods are certainly quite different; but they are very hard to assess. The question how far (22), (24), and (25) should or should not be replaced by (47), (48), and (49) depends entirely on this relative weighting, and it is difficult to feel much confidence about any particular choice. The resulting uncertainty is not great except for Δk , but is clearly not reflected in the probable errors. In the third place, it would seem at first sight logical to underweight all low-altitude declination measurements, and this includes all those made on the Sun in winter, and not only those made on sub-pole stars; at Greenwich, the winter Sun is only a few degrees higher than the limit at which sub-pole declinations are rejected altogether. The temptation to give full weight to every measurement of the Sun's declination which can be obtained in the winter months, even though the altitude is low, is a strong one, since for several months on end no observations except low-altitude ones can be obtained at all; but if the full weight is not really there it would be better to accept the fact. The weights at low altitudes can be adequately assessed for night-time observations, both in the north and in the south, and it is improbable that the weight at noon could be better. Against this must be set, however, the possibility that noon observations in summer may be so much worse than those on high-altitude stars as to be no better than the winter ones; it is the scatter in the solar residuals themselves which should really decide their formal weights. The same arguments apply in R.A. too, in principle, though the actual variation of weight may be less here. Whenever it seems even reasonably correct, taking everything into account, to assign equal weights to normal points that are equally spaced in α , the resulting simplification of the analysis is so great that it should probably be done, but in some cases it might clearly be improper.

However this may be, there can be little question that so long as the real form of the flexure function is unknown it will be impossible to eliminate most of the uncertainties discussed above, either in the solar corrections Δk and $\Delta\epsilon$, or in the systematic declination corrections of the star places which may be obtained with the instrument in question; a detailed and reliable determination of the complete flexure function would clearly still be worth a considerable effort, if it could be achieved.

The “pentag” method of determining flexure.—It seems possible that this might be done, more completely than it has in the past, by making use of an optical square, or “pentag”. This consists in principle of two plane mirrors very rigidly mounted together so as to enclose an angle of 45° ; it has the property that a ray of light incident on one mirror, in the plane perpendicular to the line joining them, and then reflected in the other mirror also, is deviated through 90° whatever its angle of incidence. If the angle between the mirrors is $45^\circ + v$, the ray is deviated by $90^\circ - 2v$. For our present purpose, the exact value of v is quite unimportant (so long as it does not exceed $1'$ or so), but its constancy must be guaranteed within a small fraction of a second. Small “squares”, as used in range-finders etc., are commonly made as solid prisms, and the most economical form is then pentagonal. When larger ones are required, fabrication from plates is more suitable; there are obviously several ways of constructing a “hollow pentag” out of fused quartz only, and it should be quite practicable to achieve the necessary permanence even with a fairly large effective aperture.

If such a pentag is mounted in front of the objective of a transit circle, by a counterpoised system supported only on the east-west axis, or on the ball-races through which the main counterpoise system acts, the circle reading for any star will always be $90^\circ - 2v$ greater, or less, than that for a direct observation of the same star, except for the difference of the two flexures. If c_1 is the reading with the pentag, and c the reading without, and if the upper sign refers throughout to the arrangement in which $c_1 \sim c + \pi/2$, we have

$$z = c + f(c) \quad (51)$$

$$z = c_1 \mp \pi/2 \pm 2v + f(c_1) \quad (52)$$

and so

$$0 = c_1 - c \mp \pi/2 \pm 2v + f(c \pm \pi/2) - f(c). \quad (53)$$

Writing $f(c) = \Sigma(a_n \cos nc + b_n \sin nc)$, we obtain the equation of condition as

$$f(c) - f(c \pm \pi/2) \mp 2v \equiv \pm 2\Sigma[\sin n\pi/4 \{a_n \sin n(c \pm \pi/4) - b_n \cos n(c \pm \pi/4)\}] \mp 2v = c_1 - c \mp \pi/2; \quad (54)$$

or, putting in successive values of n and transforming so as to remove half the double signs

$$\begin{aligned} & \sqrt{2}a_1 \cos(c \mp \pi/4) + 2a_2 \cos 2c + \sqrt{2}a_3 \cos(3c \pm \pi/4) + \dots \\ & + \sqrt{2}b_1 \sin(c \mp \pi/4) + 2b_2 \sin 2c + \sqrt{2}b_3 \sin(3c \pm \pi/4) + \dots \mp 2v \\ & = c_1 - c \mp \pi/2. \end{aligned} \quad (55)$$

The coefficients vanish identically for harmonics of order $n = 4m$, so that terms of these orders in the flexure function cannot be determined by this method. Although one may not feel entirely certain that a fourth harmonic will be negligible merely because the third and fifth ones may be found to be so, it appears somewhat improbable that any harmonic (of c) as high as the fourth can really be appreciable, unless it is used to describe a discontinuity. The only likely discontinuity would seem to be in the zenith, and a discontinuity there is readily discoverable and can with advantage be explicitly removed before proceeding to any Fourier analysis; the indeterminacy of the fourth harmonic should then be unimportant. We must, however, consider the case $n = 0$.

A zero harmonic in the “flexure” is the same thing as a constant error in all the circle readings used in flexure work or on stars; it may be put down as a

systematic error in taking the nadir, or as a "wrong" division error of the nadir division. As is well known, this is eliminated by reversal, if it really is a wrong division error; if it is a systematic error in taking the nadir which is not reversed by reversing the instrument, it is in effect a wrong co-latitude. In either case, it seems substantially irrelevant, so far as the astronomical results are concerned.

As there is no absolute term in the trigonometrical series in (55), the quantity v , the intrinsic error of manufacture in the pentag, can be obtained as one of the unknowns. Formally, it might always be convenient to do this; but it seems improbable that its value could be obtained in this way with as much accuracy as should be attainable in a pure laboratory determination. If not, and if the least-squares value always agreed with the laboratory one within their probable errors, one would of course gain by substituting the laboratory value, and transferring the term $2v$ in (55) to the right-hand side.

The unknowns in (55) would evidently all separate completely if observations of stars could be made all round the circle. The total range of c_1 -values is much greater than the range of z -values of observable stars, since the pentag allows stars to be observed when the telescope is pointing very considerably below the horizon; but the total range of c -values is of course only the same as the range of observable z , say from -80° to $+80^\circ$, and the range over which three different telescope positions are possible for the same star is appreciably more restricted again, since when the telescope is too far below the horizontal the incoming beam will foul a collimator. There is, however, no star which cannot be observed in at least two positions, and the separation of the unknowns should be quite adequate, for a moderate number of terms.

There seems no need to construct a pentag large enough to cover the entire aperture of the instrument. A four-inch beam will allow observations on a large enough number of stars for a flexure programme, and one would then presumably stop the beam down to this figure whether or not the pentag was in. The various coefficients in (55) can be evaluated for each star in the programme once for all.

If the indeterminacy of the fourth harmonic is considered undesirable, a "pentag" giving some other deflection than 90° could be constructed. The modifications in the analysis are straightforward, but the laboratory determination of the exact angle might be difficult and the astronomical one might have to be accepted; there are also some other disadvantages.

Although errors in positioning the pentag which affect the angle of incidence are completely immaterial, errors which tilt the plane of incidence are not. A small rotation from the correct position, either about the direction of the incoming beam or about the axis of the telescope, will mean that the star is observed somewhat before or after transit, and therefore not quite at culmination. The error should however be easy to allow for, since the error in transit time (which will be known) is of the first order while that in zenith distance is of the second.

It thus appears that use of a pentag should permit a better determination of the flexure than has as yet been possible by any method. It is superior to the mercury pool in permitting a second harmonic to be detected; in permitting observations over a much greater range of z -values, including all values near the zenith itself; and also in being far less liable to be interfered with by a breeze. (Since the mercury pool demands extreme steadiness in the air surrounding it, it introduces a strong temptation to reduce the pavilion aperture dangerously at all times.) The pentag is superior to all "pure" instrumental devices (even

if these were, in their own field, free from objection) in using stellar observations exclusively, so that no systematic error due to the rather different nature of a collimation, or auto-collimation, observation can enter; in addition, of course, it is superior in that it includes the effects of any circle flexure (which no pure instrumental method except the collimator one has yet succeeded in doing), and especially in being almost independent of the precision with which the auxiliary apparatus itself is mounted.

Royal Observatory,
Greenwich.

1955 March 17.

References and Notes

- (1) F. W. Bessel, *A.N.*, **3**, 209, 1825.
- (2) This has been done in the new Washington 7-inch, and it seems to have much to recommend it.
- (3) *Pub. U.S. Nav. Obs.*, XV, Part V, 150, 1948.
- (4) *Pub. U.S. Nav. Obs.*, IX, Part I, p. A CII, 1920.
- (5) M. Loewy, *C.R.*, **87**, 889, 1878; **91**, 6, 1880.
- (6) I. Bonsdorff, *Pulk. Obs. Bull.*, VI, 105, 1915.
- (7) *Greenwich Observations*, 1865, etc.; F. W. Dyson and R. T. Cullen, *M.N.*, **84**, 470, 1924.
- (8) G. B. Airy, *Mem. R.A.S.*, **32**, 9, 1864.
- (9) *Pub. U.S. Nav. Obs.*, IX, Part I, p. A CXVI, 1920.
- (10) *Pub. U.S. Nav. Obs.*, XVI, Part I, 1949.
- (11) *Pub. U.S. Nav. Obs.*, XVI, Part III, 1952.
- (12) This p.e. is taken from the previous series, as none is stated in the later work, but it should presumably hold fairly well for both.
- (13) In some cases, e.g. in *Cape Annals*, this term is written as $+\Delta\delta_0$; $\Delta\delta_0$ then has the same sign as the mean of the residuals. However, the quantities ΔL , $\Delta\epsilon$, etc., are corrections, not errors, and it seems more consistent to use the same convention for all.
- (14) *Ann. Cape Obs.*, VIII, Part IV, p. 39D, 1915.
- (15) The separability of the corrections, if the weights are equal, was noted by Gill (*Ann. Cape Obs.*, II, Part V, p. 9D, 1907), but he did not put the equation in this form.
- (16) It is improbable that the circumpolar declinations would tolerate so large a change as this, in lat. 50° ; but it has sometimes been found advisable to assume that the refraction south of the zenith is different from that to the north. If one first adds $-0''.110 \cot \phi$, i.e. $-0''.092$, and then adds $0''.110 (\cot \phi - \tan z)$ to all the solar residuals $\Delta\delta$, one has corrected the south residuals only; equations (40), (41), and (42) are in that case adjusted for a $\Delta\delta_0$ of $+0''.30 - 0''.09 = +0''.21$ and not for $\Delta\delta_0 = +0''.30$.