Automatic Data Capture and Analysis of Rayleigh Interference Fringe Displacements in Analytical Ultracentrifugation

Stephen E. Harding

Department of Applied Biochemistry & Food Science, University of Nottingham, Sutton Bonington, LE12 5RD, UK

and

Arthur J. Rowe

Department of Biochemistry, University of Leicester, Leicester LE1 7RH, UK

ABSTRACT

Low speed sedimentation equilibrium techniques using the analytical ultracentrifuge provide a powerful tool for the analysis of molecular weight and molecular weight distributions of a wide range of macro-molecular dispersions. The precision of the computed parameters depends critically upon the off-line measurement and processing of photographic images of Rayleigh interference patterns, used to record the macromolecular solute distributions at equilibrium. We have shown that it is possible to use a commercial scanning densitometer (LKB 2202), of a type widely found in laboratories in the life sciences, to capture a series of digitized scans across the fringe pattern; and have developed a simple but highly efficient algorithm, based upon the iterative evaluation of a limited number of discrete Fourier coefficients, for data analysis. This algorithm is implemented in a Pascal program which can be run on the Ultroscan's own microcomputer. Results obtained show that a gain in precision of some 5-fold has been obtained, as compared with previously fully automatic evaluation systems.

INTRODUCTION

The technique of sedimentation equilibrium in the analytical ultracentrifuge provides a powerful tool for the analysis of molecular weights,
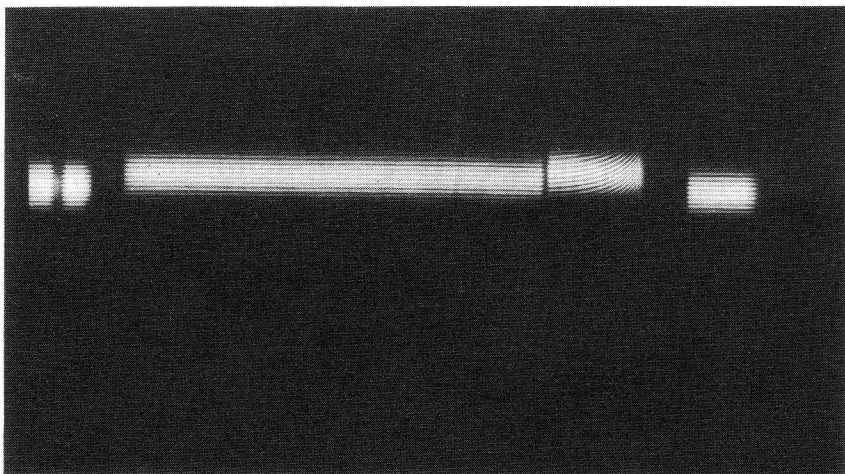
Fig. 1. Rayleigh interference fringe pattern from a low speed sedimentation equilibrium experiment on a polysaccharide (tomato pectin), using a Beckman Model E analytical ultracentrifuge. Solvent: $I = 0.10$, $pH = 6.8$. The rotor speed was 9341 rpm, the temperature 20 °C and the initial solute loading concentration was 0.4 mg ml^{-1}.

molecular weight distributions and associative interaction constants, for a wide range of solutions and dispersions of macromolecular solutes. The most accurate way of recording the distribution of the solute in the ultracentrifuge cell at sedimentation equilibrium is to photograph a Rayleigh interface optical pattern, obtained by splitting the illuminating beam between the sample sector and a sector containing solvent only.[1] The patterns thus produced (see Fig. 1 for an example) contain 7–100 parallel fringes, dependent on the system in use, and the displacement
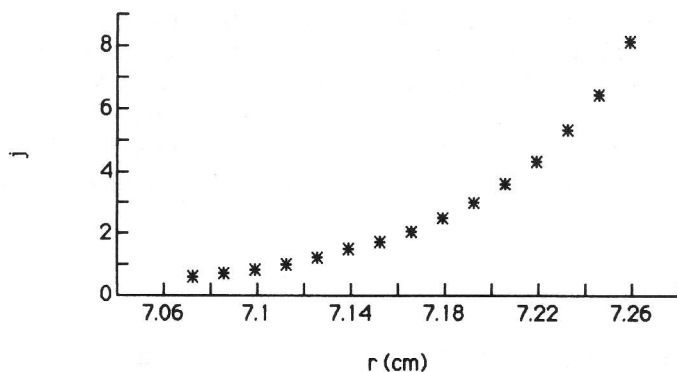
Fig. 2. Plot of fringe concentration, j, relative to that at the meniscus, as a function of radial displacement from the centre of the rotor (data from experiment in Fig. 1). Data from careful manual measurement.
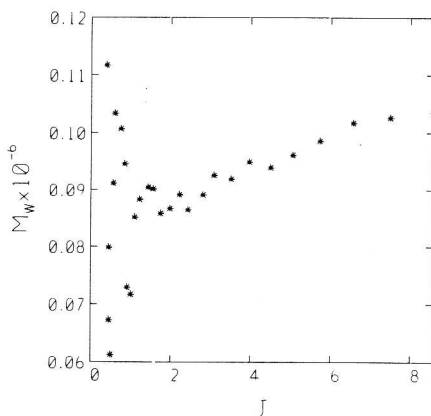
Fig. 3. Plot of the point weight average molecular weight as a function of (absolute) fringe concentration, J (data from experiment in Fig. 1).

with respect to a defined axis can be interpreted in terms of solute concentration (c) increments within the cell with respect to radial distance (r) from the centre of the rotor, on the assumption that $\partial n / \partial c =$ constant ($n =$ refractive index of the solution). The concentration increments thus measured are purely *relative,* but procedures exist for determining the absolute concentration at the meniscus of the solution column, in terms of fringes, and hence finding the absolute concentration in the same units at all other radial positions in the cell.[2]

Manual measurement of the displacement of individual fringes has been widely employed, but is both tedious and of limited precision.[3] Results from manual measurement of relative fringe increments are shown in Fig. 2. Typically a precision of some $f/50$ to $f/100$ (standard deviations) can be obtained, where f is single fringe increment. This is perfectly adequate for evaluation of molecular weights, normally from the regression of log(fringe number) upon (radius squared). The main current interest, however, concerns the analysis of thermodynamically non-ideal, polydisperse and interacting systems. This requires the use of accurate estimates for point weight (and other) average molecular weights at a range of radial positions, i.e. the data must be differenced once, or even twice. Potentially powerful mathematical techniques have been derived for the analysis of such systems,[3-5] but these and other recent developments make demands upon the precision of the basic data which cannot be met by manual methods (see Fig. 3 for an example).

The presence of multiple parallel fringes affords an obvious approach to the improvement of precision via averaging. Manual measurement then becomes even more tedious, and it is clear that for any normal use

an automatic system must be developed. Three particular difficulties arise here:

(1) the fringe pattern in most commercial ultracentrifuges is convoluted with an envelope function, arising from inhomogeneity in light intensity in the final image plane (Fig. 1);

(2) at high resolution minor imperfections in the photographic record, and even grain size and clustering, can give rise to uncertainty in fringe location;

(3) the particular application will not, in many laboratories, justify the purchase of dedicated hardware.

Automatic fringe scanning systems have been devised by several workers. By far the most comprehensive analysis of the problem has been by DeRosier *et al.*[6] They employed a two-dimensional scanning and digitising microdensitometer to scan the fringe patterns. The scans were then analysed by computation of the discrete Fourier transform of the data values; f^{-1} was found from the maximum of the modulus of the transform in the appropriate region, and the phase of the fringe scan estimated from the argument of the transform, interpolated to the region of the maximum from a small number of computed coefficients on the assumption that the contribution to the phase of the envelope function could be neglected due to its symmetry with respect to the data set. The final precision of this system was disappointing and did not exceed that obtained by manual methods for reasons which were not clear. This was also true of a simpler system built by Carlisle *et al.*,[7] which must in any case be considered theoretically questionable since it ignored the problem of the envelope function. A quite different approach was adopted by Richards and Richards.[8] They built a dedicated instrument based upon a twin photocell detector, the two photocells being arranged to be normal to the radial direction. The smoothed difference signal from the photocells becomes zero when the centre of a fringe pattern is located, and by traversing manually the relative positions of the pattern and the photocells, the logging of the signal from digital microcomparators can then be triggered automatically. A precision of $f/300$ was attained after averaging results from the five central fringes of the pattern—a very considerable improvement upon previous systems. However, even if it were to be fully automated, this approach has drawbacks. Non-standard ultracentrifuge optics have to be used to circumvent the problem of the envelope function, and dedicated apparatus has to be constructed. The results of Richards and Richards[8] do however illustrate the degree of precision which is potentially attainable.

Our own approach was to use a commercially available instrument for data capture, the LKB 2202 Ultroscan laser densitometer. This is widely used for scanning of gels and other records in laboratories in the life sciences. The use of the UCSD Pascal program GELSCAN would digitise the scans and place the values into an array of integers in a Pascal data file on disk in an Apple 11e (or IBM-compatible PC) microcomputer. Furthermore, it could also be set to record up to 50 scans at pre-specified locations ('track positions') in fully automatic mode. A photographically enlarged copy of the original interference fringe pattern may thus be digitised, and the data capture made available on disk for analysis by a suitable program written in Pascal. The manufacturer's specification for the precision and linearity of the scanning system (in the scan direction) seemed adequate for the purpose intended. The precision of the 'track positions' was not clear; we have treated this as a matter for investigation. In our context these denote radial positions in arbitrary (but easily calibrated) units.

A new algorithm for data analysis has been developed. Although Fourier analysis appears to be the obvious and natural way to analyse what is the sum of a simple harmonic function and an envelope function (with high order noise), the results obtained by this approach have been disappointing. Indeed the final precision attained by DeRosier et al.,[6] averaging a series of fringes, is no better than manual measurement of a single fringe. We have started from the recognition that what is being sought is the phase associated with a single harmonic component whose order can be pre-specified with a reasonable degree of accuracy. Rather than attempt to compute the phase as such, we have employed an iterative frame shift within the data set, searching each time for the frame shift which returns a null value for the phase, and using as a criterion for the latter the maximisation of the single Fourier coefficient of order Q, where Q is the number of fringes to be analysed, with the data set having to contain digitised values from at least $Q + 2$ fringes. It is convenient normally to use the cosine coefficient, $A(Q)$. A correction can be applied for the minimal contribution of the envelope function to the amplitude of $A(Q)$, and an interpolation carried out to determine the non-integral part of the frame shift corresponding to the maximum in $A(Q)$. The final algorithm, implemented via a UCSD Pascal program on the Apple 11e microcomputer, has been found to be stable, insensitive to rounding errors in f, and capable, in routine use on real data, of returning fringe increments with a standard deviation in the range $f/300$ to $f/500$. This precision, which is of the order one should expect to attain by any efficient algorithm analysing and averaging seven parallel fringes, opens up the possibility of applying for

the first time on a routine basis the mathematical methods referred to above.

DATA CAPTURE

The initial film record obtained from the ultracentrifuge was enlarged by approximately 15-fold onto sheet film (Agfa Rapidoprint), using the centre of the field on a Durst M800 enlarger with a Schneider–Kreuznach Componar 1 : 3·5/50 lens. Enlargement, under identical conditions, of a high-precision square-ruled grating revealed no measurable distortion of the image. The film was processed using an Agfa Rapidoprint rapid processor with attached dryer.

The enlarged image was then located on the translucent platen of the LKB Ultrascan 2202 Laser Densitometer, adjusted until the air fringes (Fig. 1) lay as accurately as possible normal to the scan direction (see 'baseline determination', below) and clipped into position. The positions of the meniscus and the start of the reference fringes, in terms of 'track numbers', were then noted as fixed datum points. The instrument was then set to scan at the desired number of enlarged radial positions, i.e. at equal intervals of track positions. Because of the geometry of the scanning spot (maximum of 800 μm width $\times$ 50 μm in the scan direction with a Gaussian intensity distribution), and also because of the limitations of disk storage space on the Apple 11e, not more than 50 radial positions within a 2–3 mm cell column were measured. A limited
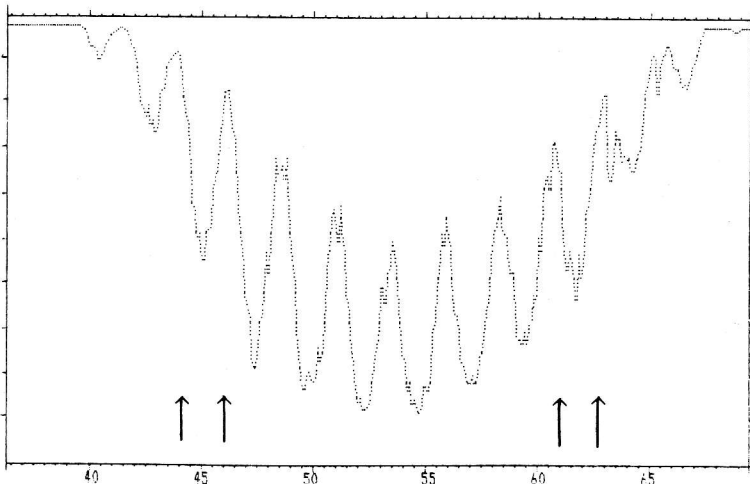
Fig. 4. Digitised optical density values output from the LKB Ultroscan, scanning at a single radial position. 1000 values are logged at each (incremented) radial position, and these form the basic data set for subsequent analysis. (Scale = 10, AbsR. = 4·0).

number of tracks were then scanned within the reference fringes to provide details of the optical baseline (below).

The GELSCAN program (LKB Instruments, Bromma, Sweden) automatically saves all the scans as DATA files on the second disk drive. A printout of a typical scan is shown in Fig. 4. The effect of the envelope function arising from the uneven light intensity distribution from the mercury arc source is very apparent; other than in the central region the true maxima/minima of the fringes cannot be specified with precision. Rather better resolution of the fringes is obtained if the resolution MODE is set to MEDIUM or to HIGH, rather than to LOW as in Fig. 4. The data analysis is then, however, correspondingly slower.

DATA ANALYSIS

The algorithm

The basis of our algorithm is the iterative maximisation of the amplitude of the Fourier cosine coefficient of order Q, where Q is the number of fringes within the data set analysed, by translation of a reading frame within the total data set. The precise degree of translation required to maximise $A(Q)$ is thus a measure of the phase shift. Throughout the present work $Q = 7$, limited by the number of clearly resolved fringes (Figs. 1 and 4), but we expect shortly to be able to analyse laser-generated fringes with $Q \geq 11$. In principle our algorithm obviously calls for the reading frame to contain an integral number of data values and to correspond to exactly Q fringes. This state can be neither recognised with certainty, given the presence of the envelope function (Figs. 1 and 4), nor attained within a precision of better than 0·5 of the translation corresponding to one data value. As shown below these conditions do not in practice limit the use of the method.

The precise sequence of logical operations is as follows:

(1) an approximate value for the number N of data values corresponding to Q fringes is estimated. In practice, using standard conditions for the optical enlargement this parameter is invariant, and is not re-measured on each occasion;

(2) a parameter *firstguess* is user-estimated for the maximum position of the first fringe;

(3) the fringe increment, f, is then estimated by iteration to maximise $A(Q)$ corrected for background amplitude (below) within the plane $(\textit{firstguess} \pm \partial) \times (N \pm \partial)$, where ∂ is a

user-supplied integral parameter, usually 1 or 2. The value of N is then refined if necessary to be the nearest integer value to $f \cdot Q$;

(4)   for each scan, the maximum value of $A(Q)$ within the reading frame starting at (*firstguess* $\pm \partial$) is found as described below, and the (non-integral) deviation of the reading frame from *firstguess* is added to *firstguess* to give the estimate for the phase shift and hence the fringe displacement;

(5)   when the cumulative fringe displacement exceeds N/Q the reading frame is reset by N/Q and the cumulative fringe displacement incremented thereafter by f (i.e. by the increment due to a single fringe).

Because of the small initial fringe increments, the initial value of *firstguess* suffices for the first 2 scans. Thereafter, this parameter is continuously refined by simple forward extrapolation; a procedure also adequately accurate to detect the possibility of more than a single fringe increment occurring between scans.

The UCSD Pascal program ANALYSER implements the above algorithm. Using efficient procedures avoiding unnecessary re-computation of cosine functions, and a quadratic fit interpolation to minimise iteration (below), the time to estimate a fringe increment at a single radial position can be kept to less than 10 seconds on the Apple 11e (with accelerator board). Hence, even a comprehensive set of scans can be analysed within a reasonable period of time.

The validity of the algorithm and the precision of the various procedures has been thoroughly studied, as shown below.

Detection of the fringe pattern by measurement of $A(Q)$

The amplitude of the cosine coefficients reflects the envelope function (in the lower orders) and random noise (in the higher order coefficients). We have found that for $Q \geq 7$ there is very little contribution of the envelope function to the coefficients in the range $Q \pm 2$, whilst the noise spectrum is very far removed. The four 'flanking coefficients' routinely approximate to zero (Fig. 5), and any small contribution to the amplitude of $A(Q)$ is removed by the procedure QUADINTERP which interpolates this contribution after fitting a least-squares quadratic function to the four flanking coefficients. The corrected $A(Q)$ can thus be considered to arise solely from the fringes.
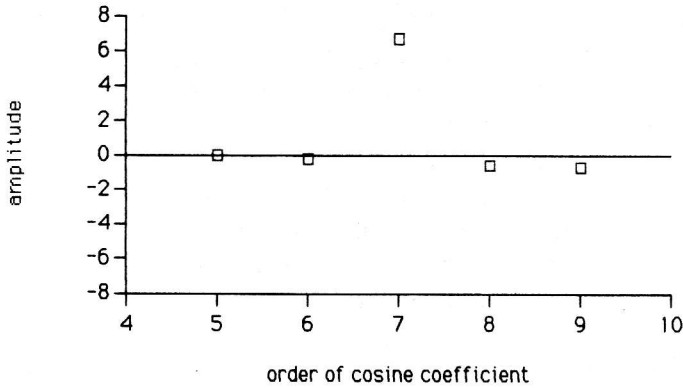
Fig. 5. The amplitude of the Fourier cosine coefficients as a function of order of coefficient, computed for a typical data set (169 central values, data shown in Fig. 4).

Determination of the fringe number, f (wavelength)

The value for the width of a single fringe, f, must be known accurately if errors in the integral count of the fringes are to be minimised. Yet since an integral and finite number of data values must be used in the frame shift procedure, an approximated value for $f(=N/Q)$ must of necessity be used. In Fig. 6 we see the effect of varying N upon the amplitude of $A(Q)$, where each $A(Q)$ is a value maximised by the frame shift procedure. It is apparent that:

(1) the amplitude of the maximal value of $A(Q)$ varies only slowly with the assumed value for N. Since N can hardly be in error by
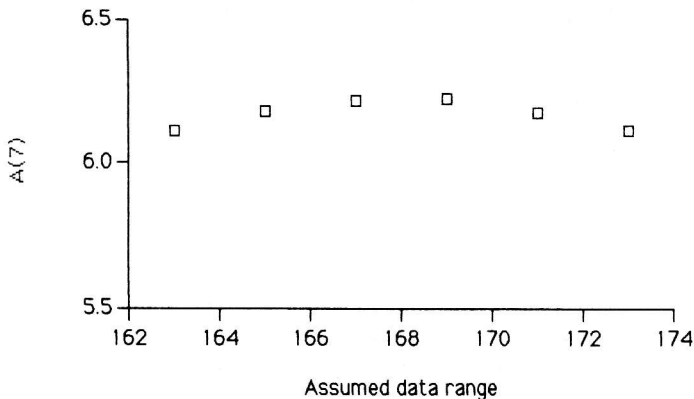
Fig. 6. The effect of varying the number of data points assumed to correspond to 7 fringes ($7f$) upon the amplitude of $A(7)$. In each case the value of the latter is the maximum obtained iteratively by frame shifting (see Fig. 7).

more than a single integer, it follows that no significant errors are likely to be introduced by the approximation involved;

(2)   the variation of maximal $A(Q)$ with assumed N is smooth, and an interpolated maximal value could be found. We used a least squares quadratic fit to the $A(Q)$ values to achieve this. With the central value designated at the zeroth in N, and fitting a total of 5 $A(Q)$ values, the interpolated value in N is found by elementary algebra from $-a_0/2$, where a_0 is the zeroth order coefficient of the quadratic fit.

Variation of $A(Q)$ with frameshift

The algorithm will only be stable if $A(Q)$ varies strongly with frame shift, and shows a clearly defined maximum. Results from real data are shown in Fig. 7 for $\partial = 2$. Clearly these conditions are met. The interpolated non-integral part of the frame shift is again found by a least squares quadratic fit to the $A(Q)$ values. For the highest precision working the program permits an iterative fit with decreasing ∂ value (3 down to 1) to be performed, but the gain is found to be marginal compared with maintaining $\partial = 2$.

Determination of baseline

By use of the Ultroscan for data capture and our algorithm for data analysis, a precision in fringe increment is attained which very significantly exceeds that obtainable by direct optical measurement. It therefore follows that the alignment of the enlarged photographic
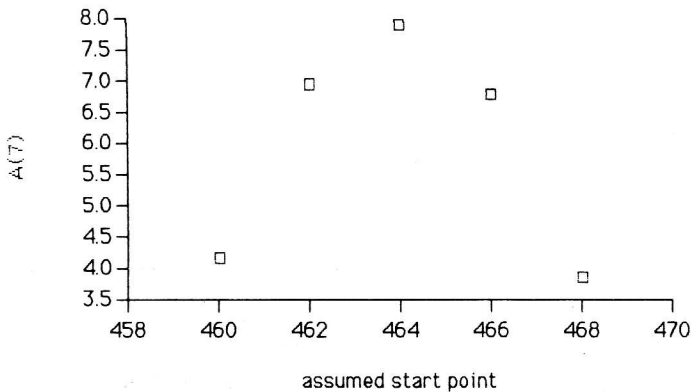
Fig. 7. Plot of the amplitude of the cosine coefficient $A(7)$ as a function of assumed start point within the data set, with an iterative frame shift $\partial = 2$ values. 169 values were taken to correspond to 7 fringes (see Fig. 6).

record in the densitometer cannot be achieved by eye with the desired accuracy. We therefore routinely measure the air or reference fringes at a series of radial positions, and by simple linear regression analysis of the measured fringe increment obtain the true optical baseline. This can then be used to correct the data obtained from the region of the solution column.

Precision of results obtained

Precision was estimated by both the point variance of the fringe increments from the linear regression line fitted to values from reference fringes, and, since the equilibrium trace (fringe increment vs radius) approximates rather closely to a quadratic function, from the variance of the second difference of the data values in a full experimental run. By either method we find standard deviations in the range $f/300$ to $f/500$.

RESULTS WITH EXPERIMENTAL DATA

Figure 8 shows a plot of fringe increment against radial position for a data set taken from a typical sedimentation equilibrium experiment. The smoothness of the final data set (see Fig. 2) is such that it is difficult to give an impression by graphical means of the residual noise. Figure 8 may, however, legitimately be compared with the optimal results obtained by DeRosier *et al*,[6] where their much poorer final precision is associated with residual noise at a level which is visually obvious (see Figs. 3 and 4[6] with Fig. 8, present work). Our final
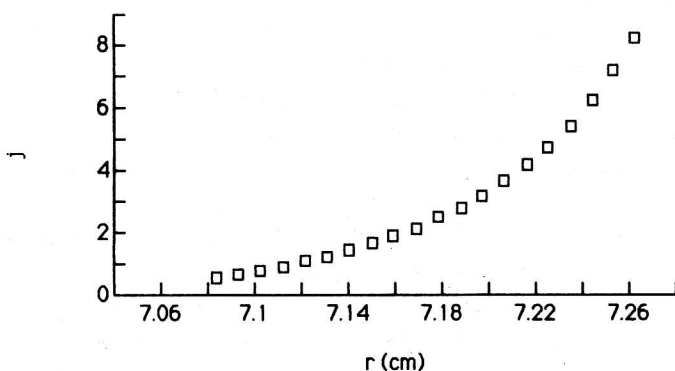
Fig. 8. Plot of estimates for relative fringe concentration, *j*, as a function of radial displacement from the centre of the rotor, yielded by the program ANALYSER applied to data logged at equal increments in *r* by the LKB Ultroscan 2202 laser densitometer.

precision is such that differencing of the data can now be carried out with some reasonable degree of smoothness remaining in the differenced values.

DISCUSSION

Our results demonstrate unequivocally that by using the LKB Ultroscan 2202 for data capture and a newly devised algorithm for data analysis, a final precision of $f/300$ to $f/500$ in measured fringe displacement is readily obtained from typical real experimental data. This equals or exceeds the precision obtained using a semi-automatic system based upon twin photocells and a light-difference detector.[8] A major initial uncertainty in our case was the lack of knowledge as to the accuracy and reproducibility of the track position settings in the ultroscan. Our finding that the standard deviation of measurements made on reference fringes (nearly level and hence insensitive to errors in track position) was similar to that found with steeply curved fringes shows that, under the conditions used at least, the instrumental capability was adequate. There is a degree of inconvenience, however, in having to enter the track positions selected into the computer manually; a two-dimensional scanning instrument would, therefore, be more convenient in use, though not necessarily more accurate.

The final precision achieved by our fringe analysis procedure, though hardly unprecedented in general terms, is considerably better than that which has previously been achieved by an automatic data capture system in the context of this particular application. As noted earlier, the system devised by DeRosier et al.[6] did not exceed $f/100$ at best, and other (less sophisticated) published systems have shown no improvement (see Introduction). DeRosier et al. speculate that their densitometer may have been prone to backlash, and clearly a hardware problem could have been at least partly responsible for the limited final precision obtained. Digital micrometers should however be capable of functioning satisfactorily at this level, and the semi-automatic system of Richards and Richards[8] achieved a precision of $f/300$ using a mechanical stage. The algorithm of DeRosier et al.[6] is not totally secure. The assumption of zero phase contribution from the envelope function in the region of the transform, where the modulus arising from the fringe pattern is maximal, is approximate only, and must become hazardous when small values of phase shift are being estimated. Earlier work in one of our laboratories (A. J. Gibbs and A. J. Rowe, unpublished data) found serious difficulties in making reliable estimates of small

phase shifts, applying an algorithm similar to that of DeRosier *et al.* to simulated data. Furthermore, the precision expected must be expected on any simple theory to be a function of the numerical value of the dependent parameter (in this case the phase) being estimated. The use of a null procedure in which final precision will be equal for all values measured is therefore desirable on general grounds.

Our algorithm therefore provides a very fast and stable way of extracting detail concerning the relative phase shift of a single harmonic component, by computing only five individual Fourier (cosine) coefficients for a small number (usually five) of frame shifts corresponding to the same number of assumed fringe increments, and by avoiding further iterations by means of a numerical evaluation of the maximal function value. This algorithm may not be without use in other fringe analysis contexts. In our own field of application it offers the possibility that a series of most interesting methods derived during the past 20 years,[3-5,9,10] which have much potential for protein chemistry but which have proved difficult to implement because of the severe demands placed upon smoothness of data, may now turn out to be usable; given that the whole of the data capture and evaluation can be performed in fully automatic mode using an instrument already commonly found in laboratories in the life sciences.

REFERENCES

1. P. H. Lloyd, *Optical Methods in Ultracentrifugation, Electrophoresis and Diffusion*, Oxford University Press, Oxford, 1974.
2. J. M. Creeth and S. E. Harding, Some observations on a new type of point average molecular weight, *J. Biochem. Biophys. Meth.*, **7** (1982) 25–34.
3. D. C. Teller, Characterisation of proteins by sedimentation equilibrium in the analytical ultracentrifuge, *Methods in Enzymology*, **27** (1973) 371–90.
4. D. C. Teller, T. A. Horbett, E. G. Richards and H. K. Schachman, Ultracentrifuge studies with Rayleigh interference optics III: computational methods applied to high speed sedimentation equilibrium experiments, *Ann. NY Acad. Sci. (USA)*, **164** (1969) 66–101.
5. D. E. Roark and D. A. Yphantis, Studies of self-associating systems by equilibrium centrifugation, *Ann. NY Acad. Sci. (USA)*, **164** (1969) 245–78.
6. D. J. DeRosier, P. Munk and D. J. Cox, Automatic measurement of interference fringe photographs from the ultracentrifuge, *Anal Biochem.*, **50** (1972) 139–53.
7. R. M. Carlisle, J. I. H. Patterson and D. E. Roark, An automatic microcomparator for ultracentrifuge fringe measurements, *Anal. Biochem*, **61** (1974) 248–63.
8. E. G. Richards and J. H. Richards, Light-difference detector for reading

Rayleigh fringe patterns from the ultracentrifuge, *Anal. Biochem.*, **62** (1974) 523–30.

9. S. E. Harding, The representation of equilibrium solute distributions for non-ideal polydisperse systems in the analytical ultracentrifuge: application to mucus glycoproteins, *Biophys. J.*, **47** (1985) 247–50.

10. H. J. Rowe and A. J. Rowe, A re-examination of the equations appertaining to sedimentation equilibrium, *Biochem. Biophys. Acta*, **222** (1970) 647–59.