Direct Sedimentation Analysis of Interference Optical Data in Analytical Ultracentrifugation

Peter Schuck* and Borries Demeler[#]

*Molecular Interactions Resource, Bioengineering and Physical Science Program, ORS, National Institutes of Health, Bethesda, Maryland 20892, and [#]Department of Biochemistry, University of Texas Health Sciences Center at San Antonio, San Antonio, Texas 78284 USA

ABSTRACT   Sedimentation data acquired with the interference optical scanning system of the Optima XL-I analytical ultracentrifuge can exhibit time-invariant noise components, as well as small radial-invariant baseline offsets, both superimposed onto the radial fringe shift data resulting from the macromolecular solute distribution. A well-established method for the interpretation of such ultracentrifugation data is based on the analysis of time-differences of the measured fringe profiles, such as employed in the g(s*) method. We demonstrate how the technique of separation of linear and nonlinear parameters can be used in the modeling of interference data by unraveling the time-invariant and radial-invariant noise components. This allows the direct application of the recently developed approximate analytical and numerical solutions of the Lamm equation to the analysis of interference optical fringe profiles. The presented method is statistically advantageous since it does not require the differentiation of the data and the model functions. The method is demonstrated on experimental data and compared with the results of a g(s*) analysis. It is also demonstrated that the calculation of time-invariant noise components can be useful in the analysis of absorbance optical data. They can be extracted from data acquired during the approach to equilibrium, and can be used to increase the reliability of the results obtained from a sedimentation equilibrium analysis.

INTRODUCTION

In recent years, new ultracentrifugal methods for modeling the time course of sedimentation of macromolecules have received much attention (Correia, 1998; Laue, 1997; Schuster and Toedt, 1996; Stafford, 1997). In particular, techniques that directly utilize solutions of the Lamm equation (Fujita, 1962; Lamm, 1929) for data analysis are under rapid development. This includes both improved approximate analytical solutions (Behlke and Ristau, 1997; Holladay, 1979; Philo, 1997), as well as new approaches for the efficient numerical solution of the Lamm equation (Demeler and Saber, 1998; Schuck, 1998; Schuck et al., 1998; Stafford, 1998). These new sedimentation velocity methods are attractive, for example, because of their potential for the rigorous analysis of sedimentation profiles of small peptides (Philo, 1997; Schuck et al., 1998) and of interacting systems of macromolecules (Bethune and Kegeles, 1961a, b; Cann and Kegeles, 1974; Claverie, 1976; Cox, 1969; Cox and Dale, 1981; Gilbert and Gilbert, 1973; Goad and Cann, 1969; Schuck, 1998; Stafford, 1998), and they are complemented by improved methods for hydrodynamic modeling and for the interpretation of the sedimentation coefficients of macromolecules (Bloomfield et al., 1967; Byron, 1997; de la Torre, 1992). However, these methods suffer from the disadvantage that they cannot be directly applied to the analysis of analytical ultracentrifugation data acquired with the interference optical system.

The interference data acquisition system, as developed by Yphantis, Laue, and co-workers (Laue, 1994; Laue et al., 1983, 1984; Yphantis et al., 1984, 1994), has several virtues that can make it the method of choice for a number of different types of ultracentrifuge experiments. Refractometric detection is currently the only optical method that can be applied if the solutes under investigation have no absorption, such as many synthetic polymers or polysaccharides. Interference optical detection can also be very helpful if the solutes under investigation have extinction coefficients that are too small or too large, respectively, to be used at practical concentrations in the absorption optical system, or if the absorbance of the solute of interest is superimposed by a large absorbance of buffer components. For example, ATP or GTP binding proteins may require significant amounts of nucleotide in the buffer during an experiment, rendering absorbance analysis of such proteins next to impossible. In the Optima XL-I, the linearity of the signal and the precision and the sensitivity for proteins are very high, and can exceed those of the absorbance optics. This is true, in particular, if buffer components (such as HEPES and DTT) do not permit the exploitation of the protein backbone extinction in the far UV. In addition, since interference patterns are recorded from the entire solution column at once, data acquisition is much more rapid. This can increase the accuracy of the recorded time course, and permits the accumulation of considerably larger data sets, leading to improved confidence limits in the sought parameters.

Unfortunately, interference optical data contain special noise components that make them difficult to interpret directly. In addition to random noise, interference data contain large, systematic time-invariant noise components (abbreviated in the following as TI), as well as significant systematic radial-invariant (RI) noise components. Both types

of error are superimposed onto the interference fringe shifts arising from the macromolecular solute distribution. The elimination of the TI noise by time-derivative methods (*dc/dt*) and g(s*) transformations have proven to be most useful (Stafford, 1992; Yphantis, 1984). The g(s*) method is based on a special case of the Lamm equation for ideal noninteracting macromolecules in an infinitely long solution column. This is a good approximation for larger macromolecules and higher rotor speeds, where the meniscus can be depleted, and where negligible back-diffusion from the bottom of the cell allows clear plateaus to be obtained. Unfortunately, these conditions do not allow direct application to small molecules or systems of interacting solutes.

The novel numerical methods for the hydrodynamic shape analysis of small molecules and the methods employing numerical solutions of the Lamm equation can, in principle, be adapted to the time-derivative analysis through their extension to the calculation of differences $c(t + \Delta t) - c(t)$. These can then be used in the modeling of experimental time-difference data (Stafford, 1998; J. S. Philo, personal communication). Although this approach does allow the modeling of interference optical ultracentrifuge data, for statistical and computational reasons it does not seem completely satisfying. First, it requires the calculation of the time differences of noisy data, and if smoothing or averaging is applied to the data, this may introduce a small bias (Johnson and Straume, 1994). Second, the time-derivative transformation does not permit the well-known statistical tests of goodness-of-fit that are related to the direct modeling of raw experimental data (Beechem, 1992; Johnson and Straume, 1994). The goal of the present paper is the presentation of an alternative approach allowing the direct least-squares modeling of interference optical data in a way that takes into consideration the special noise structure. Furthermore, we demonstrate its applicability and an analysis of the statistical accuracy of the derived parameters.

THEORY

Direct fitting of the time-invariant noise

The problem of finding the least-squares fit of experimental data, $a_n(r_i)$, by modeling with a superposition of solutions of the Lamm equation, $\mathbf{L}(t_n, r_i, s_k, D_k)$, and an unknown time-invariant (TI) noise component $b(r_i)$ can be written as

$$\underset{n_k, D_k, c_k, b_i}{\text{Min}} \sum_n \sum_i \left[a_{n,i} - \left(b_i + \sum_k c_k \mathbf{L}_{n,i}^{(k)} \right) \right]^2 \quad (1)$$

where $a_{n,i} = a_n(r_i)$ denotes the experimental data of scan n, recorded at time t_n and at radius r_i, $\mathbf{L}_{n,i}^{(k)} = \mathbf{L}(t_n, r_i, s_k, D_k)$ denotes the solution of the Lamm equation for a component k with sedimentation and diffusion coefficient s_k and D_k, respectively, c_k denotes the loading concentration of species k, and $b_i = b(r_i)$ denotes the TI noise, or equivalently, a local baseline offset at each radial grid point of the scan. The summation is over all N scans, with R radial data points

per scan included in the analysis, which gives a total of $N_{tot} = N \times R$ data points.

Because of the excessive number of unknowns that are introduced by the unknown baseline components $b_i = b(r_i)$ (>1000 for a long solution column), this least-squares problems may seem to be extremely difficult to solve. However, these unknowns can be directly eliminated by following the strategy of separation of linear and nonlinear parameters (Ruhe and Wedin, 1980). It permits the algebraical calculation of the least-squares values of the linear parameters b_i and c_k for any given set of nonlinear parameters, i.e., Lamm equation solutions $\mathbf{L}(t_n, r_i, s_k, D_k)$.

By requiring the partial derivatives of Eq. 1 with respect to the baseline parameters b_j to vanish, we get

$$0 = \sum_n \sum_i \left[a_{n,i} - \left(b_i + \sum_k c_k \mathbf{L}_{n,i}^{(k)} \right) \right] \delta_{ij} \quad \text{for all } j$$

$$b_j = \bar{a}_j - \sum_k c_k \bar{\mathbf{L}}_j^{(k)}$$

$$(2)$$

where

$$\bar{a}_j = \sum_n a_{n,j}/N_s \quad \text{and} \quad \bar{\mathbf{L}}_j^{(k)} = \sum_n \mathbf{L}_{n,j}^{(k)}/N$$

are the average signal and Lamm equation solution, respectively, at radius r_j. Similarly, setting the partial derivatives of Eq. 1 with respect to the concentrations c_l zero and inserting Eq. 2, we can obtain a system of linear equations for the concentrations c_k:

$$0 = \sum_n \sum_i \left[a_{n,i} - \left(\bar{a}_i - \sum_k c_k \bar{\mathbf{L}}_i^{(k)} \right. \right.$$

$$\left. \left. + \sum_k c_k \mathbf{L}_{n,i}^{(k)} \right) \right] (\mathbf{L}_{n,i}^{(l)} - \bar{\mathbf{L}}_i^{(l)}) \quad \text{for all } l$$

$$\sum_n \sum_i (a_{n,i} - \bar{a}_i)(\mathbf{L}_{n,i}^{(l)} - \bar{\mathbf{L}}_i^{(l)})$$

$$= \sum_k c_k \left[\sum_n \sum_i (\bar{\mathbf{L}}_i^{(k)} - \mathbf{L}_{n,i}^{(k)})(\mathbf{L}_{n,i}^{(l)} - \bar{\mathbf{L}}_i^{(l)}) \right] \quad (3)$$

These algebraic Eqs. 2 and 3 allow direct calculation of the least-squares concentrations c_k and all TI noise components b_i (positivity of the concentrations can easily be achieved using the algebraic methods described in Lawson and Hanson, 1974). Even though the number of parameters b_i to be determined can be very large, this algebraic calculation is computationally very inexpensive compared with the numeric solution of the Lamm equation.

Therefore, it is advantageous to solve the minimization problem by nonlinear regression for the nonlinear parameters s_k and D_k as unknowns, using conventional curve-

fitting routines,

$$\underset{s_k, D_k}{\text{Min}} \sum_n \sum_i [a_{n,i} - \mathbf{L}_{n,i}^{*(k)}(s_k, D_k)]^2 \qquad (4)$$

while at the same time algebraically calculating for each set of (s_k, D_k) the least-squares values $c_k(s_k, D_k)$ and $b_i(s_k, D_k)$ and the best-fit linear combination

$$\mathbf{L}_{n,i}^{*(k)}(s_k, D_k) = b_i(s_k, D_k) + \sum_k c_k(s_k, D_k)\mathbf{L}_{n,i}^{(k)}(s_k, D_k), \quad (5)$$

through Eqs. 2 and 3. As will be described in the following, an analogous strategy can be used for the case of a baseline offset that is constant in time and throughout the solution column.

Direct fitting of the time-invariant and radial-invariant noise

Due to imperfections in the data acquisition system (which may be caused by thermal or mechanical instabilities), it may be necessary to analyze a set of experimental data allowing each of the scans to have a different, small, and unknown radial-invariant (RI) noise component β_n. The modeling with these constant radial baseline offsets β_n local to each scan, in addition to the TI noise b_i common to all scans (but changing with radius), requires the solution of:

$$\underset{s_k, D_k, c_k, b_i \beta_n}{\text{Min}} \sum_n \sum_i \left[a_{n,i} - \left(\beta_n + b_i + \sum_k c_k \mathbf{L}_{n,i}^{(k)} \right) \right]^2 \qquad (6)$$

It is obvious that the two perpendicular baseline contributions are linearly dependent. We can eliminate this problem by constraining the TI noise b_i to have a vanishing mean value.

$$\sum_i b_i = 0 \qquad (7)$$

Again, all the linear parameters β_n, b_i, and c_k can be algebraically optimized by solution of the linear equation system obtained for vanishing partial derivatives. Setting the partial derivatives with respect to the RI noise β_n zero and using Eq. 7 leads to

$$\beta_n = \frac{1}{R_r} \sum_i \left(a_{n,i} - b_i - \sum_k c_k \mathbf{L}_{n,i}^{(k)} \right)$$

$$= \frac{1}{R_r} \sum_i a_{n,i} - \sum_k c_k \frac{1}{R_r} \sum_i \mathbf{L}_{n,i}^{(k)}$$

$$=: \tilde{a}_n - \sum_k c_k \tilde{\mathbf{L}}_n^{(k)} \qquad (8)$$

Insertion of this result into Eq. 6 leads to a new problem

$$\underset{s_k, D_k, c_k, b_i}{\text{Min}} \sum_n \sum_i \left[\left(a_{n,i} - \tilde{a}_n \right) - \left\{ b_i + \sum_k c_k(\mathbf{L}_{n,i}^{(k)} - \tilde{\mathbf{L}}_n^{(k)}) \right\} \right]^2 \qquad (9)$$

which is of the same form as Eq. 1, and can be directly solved for the remaining linear parameters b_i and c_k using the strategy outlined above. This, in turn, will allow the application of Eq. 8 for the calculation of the RI noise components $\beta_n(s_k, D_k)$.

EXPERIMENTAL

Sedimentation velocity experiments were performed with a Beckman Optima XL-I analytical ultracentrifuge equipped with an interference optical system, and with a Beckman Optima XL-A analytical ultracentrifuge with absorbance optics, respectively. Myoglobin and gamma globulin samples (purchased from Sigma, St. Louis, MO) were redissolved in PBS, and epon double-sector centerpieces equipped with sapphire windows were filled with 350 μl of sample PBS, respectively. Using an An50-Ti rotor, the samples were centrifuged at rotor speeds of 40,000 rpm at a temperature of 25°C. Two independent sedimentation velocity experiments were performed, one in the XL-I with interference patterns acquired in time intervals of 1 min, and one in the XL-A using higher loading concentrations, with radial absorbance distributions scanned in intervals of 12 min.

Lamm equation analysis was performed using *sedfit*, a sedimentation velocity data analysis software that uses principles for numerical solutions described in Schuck (1998) and Schuck et al. (1998). This Lamm equation analysis software is available on request, or can be downloaded from http://www.biochem.uthscsa.edu/auc/software, or from ftp://rasmb.bbri. org/rasmb/spin. The Svedberg equation was employed for the calculation of diffusion coefficients based on the buoyant molar mass and the sedimentation coefficients of the solutes (Svedberg and Pedersen, 1940). To account for the time needed for acceleration of the rotor, the time after start of the centrifuge for each scan was calculated from the $\omega^2 t$ entry of each file. Exact location of the meniscus and the bottom of the solution column were treated as fitting parameters to be optimized after manual graphical predetermination. Integral fringe shifts were eliminated before data analysis. Problems due to the slightly varying radial increments in the acquisition of the absorbance data were eliminated for the calculation of the systematic noise components by expansion of each data set via interpolation to a larger grid that includes the experimentally acquired points of all scans. This produces a set of scans with a common radial grid; least-squares analysis, however, was performed with the original data sets. All error estimates include cross-correlation of all the parameters, and were based on F-statistics (Bevington and Robinson, 1992; Press et al., 1992). g(s*) analysis was performed with the program *dcdt 30*, kindly provided by Dr. Walter Stafford, Boston Biomedical Research Institute.

RESULTS

Fig. 1 shows a set of experimental interference data acquired during the sedimentation of myoglobin. It is obvious that these data exhibit a time-invariant background signal (or TI noise) with a magnitude that is comparable to the signal contribution from the sedimentation of the macromolecule. The acquired 200 scans were decomposed into the TI and RI noise components, and theoretical sedimentation profiles based on the Lamm equation for an ideal single species with the molar mass of myoglobin. The calculation of the linear parameters, i.e., the orthogonal noise components and the solute loading concentration, is
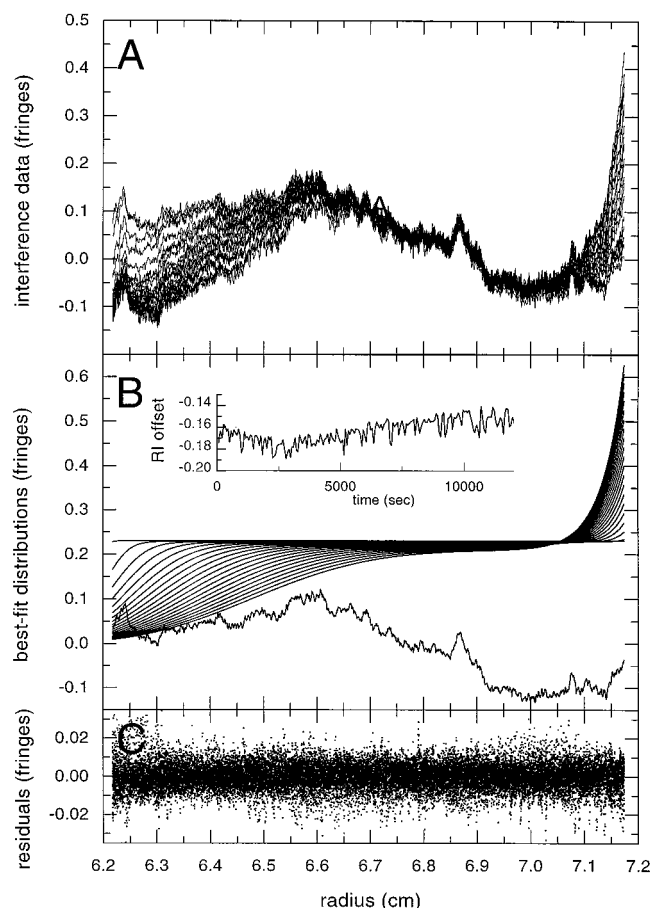
FIGURE 1 Analysis of the interference data of the sedimentation myoglobin. (A) Experimental fringe displacement profile. For clarity, only every 10th scan of the total of 200 is shown. (B) Best-fit decomposition into the calculated sedimentation profiles of a single ideal component (constraining the molar mass to 17200 Da and a partial specific volume of 0.75 ml/g), and into the TI and RI noise components. The inset shows the RI noise contributions as a function of time. The calculated TI noise is shown below the calculated solutions of the Lamm equation. Best-fit sedimentation coefficient from analysis of all 200 scans is $s_{myo} = 2.22$ S, with a loading concentration of 0.237 fringes. (C) Residuals of the fit, which has a rmsd of 0.0065 fringes in the complete analysis of the 200 scans.
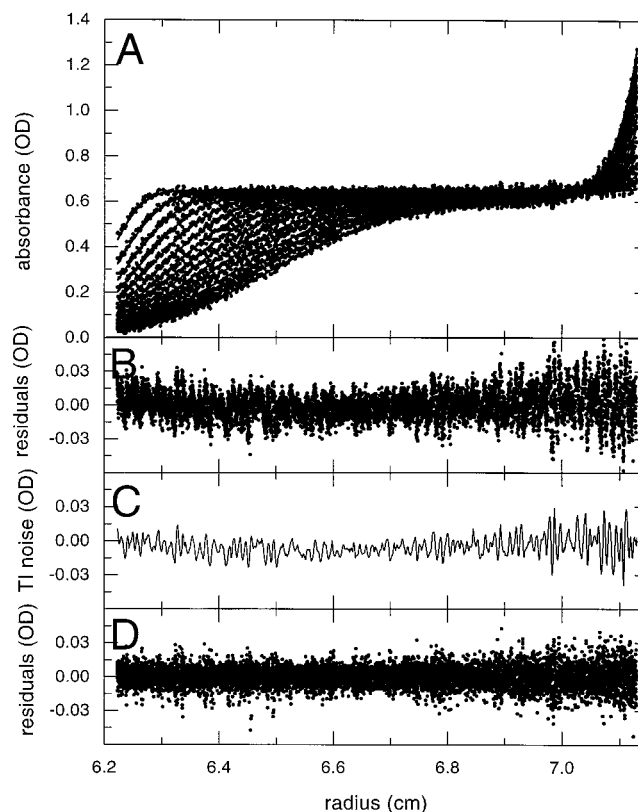

FIGURE 2 Absorbance profiles of the sedimentation of myoglobin. The absorbance distributions were scanned at 465 nm, in time intervals of 12 min for a total experimental time approximately identical to the data shown in Fig. 1. (A) Experimental (circles) and best-fit distributions (lines) constraining the sedimentation coefficient to $s_{myo} = 2.27$ S, as obtained in the analysis of the data shown in Fig. 1. (B) Residuals of the fit with a rmsd of 0.0127 OD. (C) TI noise component, calculated via Eq. 2 on the basis of the same model as (A) and (B). (D) Residuals of the extended model, rmsd = 0.0088 OD.

computationally very inexpensive. Minimization with respect to the sedimentation coefficient and the systematic noise components leads to a well-determined best-fit sedimentation coefficient of $s_{myo} = 2.22$ S. It is apparent in Fig. 1 that this model describes the data very well, with a rms deviation of 0.0065 fringes. The error estimate σ_{myo} for the analysis of 200 scans is $\ll 0.05$ S; for the analysis of only the 20 data sets shown in Fig. 1, the error estimate σ_{myo} is 0.1 S. It is also apparent from the calculated RI noise components, and from inspection of the experimental data at the hinge point (at ~ 7.05 cm), that this time-dependent vertical displacement of the scans is significant.

For comparison, the same sample at a higher concentration was used in a sedimentation velocity experiment with an absorbance optical scanner. Fig. 2 shows the absorbance distributions obtained at a wavelength of 465 nm within

approximately the same experimental time interval. Modeling of these data using the calculated sedimentation coefficient from the previous interference experiment and performing a least-squares optimization of the loading absorbance and baseline offset leads to a rms deviation of 0.0127 OD (Fig. 2). This demonstrates that the direct analysis of the interference profiles solving for the unknown TI and RI noise components is consistent with conventional analysis of absorbance distributions. The best-fit sedimentation coefficient of these absorbance data is 2.27 ± 0.04 S, with a final rmsd of 0.0125 OD. Within the precision σ_{myo} imposed by instrumental error, both best-fit results from the interference and from the absorbance experiments are identical.

For the assessment of the correlation of the systematic noise parameters with the sedimentation coefficient, we allowed next for a TI noise component in the analysis of the absorbance data. The value of this exercise is the ability to compare a data analysis with and without consideration of TI and RI noise components under conditions where both approaches would be reasonable assumptions. It is noteworthy that under consideration of a TI noise component in the

otherwise unchanged model of Fig. 2, the rms deviation to the experimental data drops from a value of 0.0125 OD to 0.0088 OD (Fig. 2 *D*). This indicates that small window imperfections can contribute to a time-invariant absorbance signal, which is apparent also in the distribution of the residuals of the absorbance analysis (Fig. 2 *B*). When TI noise is considered, the best-fit sedimentation coefficient is 2.23 ± 0.05 S. The only small increase of the error estimate σ_{myo} from 0.04 S to 0.05 S suggests that the calculation of the TI noise components does not substantially decrease the precision of the calculated sedimentation coefficient. A similar result is obtained if possible RI noise components are taken into account ($s = 2.29 ± 0.05$ S). However, the precision of the sedimentation coefficient decreases to a value σ_{myo} of 0.08 S, near the 0.1 S observed in the analysis of the interference data, if both TI and RI noise is considered in the model ($s = 2.24 ± 0.08$ S).

As an alternative treatment of the RI noise, we aligned (in a least-squares sense) the interference patterns in the air-to-air region above the solution column, where the interference patterns are not influenced by any sedimentation processes in the solution. Since in this region the RI noise components are still visible, this treatment of the data should in principle eliminate the need for taking RI noise into account in modeling the data. The magnitude of the calculated RI noise contributions and the air-to-air aligned jitter is shown in Fig. 3. While both methods coincide very well in the description of the slow drifts of the baseline offset, it is apparent from Fig. 3 that they differ in their high-frequency components, which are of the same magnitude as the slower drifts. After air-to-air alignment, the best-fit sedimentation coefficient from the analysis of the 200 scans without further consideration of RI noise was 2.23 S, with a rms deviation of 0.0080 fringes. This magnitude of the rms deviation is lower than the value obtained in the analysis without any consideration of RI noise (0.0111), but still significantly higher then the rms error obtained with the algebraic calculation of RI noise (0.0065). This suggests that the air-to-air alignment is not completely effective. Furthermore, the statistical analysis of the air-to-air aligned subset of the 20 interference data sets as shown in Fig. 1 leads to a considerably lower *s*-value of 2.06 S, and an only slightly reduced error estimate of $\sigma_{myo} = 0.09$ S.

For comparison of the presented noise decomposition with the results of a g(s*) analysis (Stafford, 1992), Fig. 4 shows the interference profiles of the sedimentation of gamma globulin. This larger molecule produces sedimentation profiles that exhibit a distinct sedimentation boundary, which includes a region of depletion near the meniscus as well as a stable plateau at the bottom of the cell, thereby allowing the application of the g(s*) method for data transformation. The g(s*) data transformation displays a slightly asymmetrical peak with a maximum at a smaller *s*-value (Fig. 4). This profile suggests the presence of a second component. A Lamm analysis with a model including potential aggregation of the gamma globulin into a dimer yields *s*-values of $s_{\gamma 1} = 7.72$ S and $s_{\gamma 2} = 11.6$ S, with a
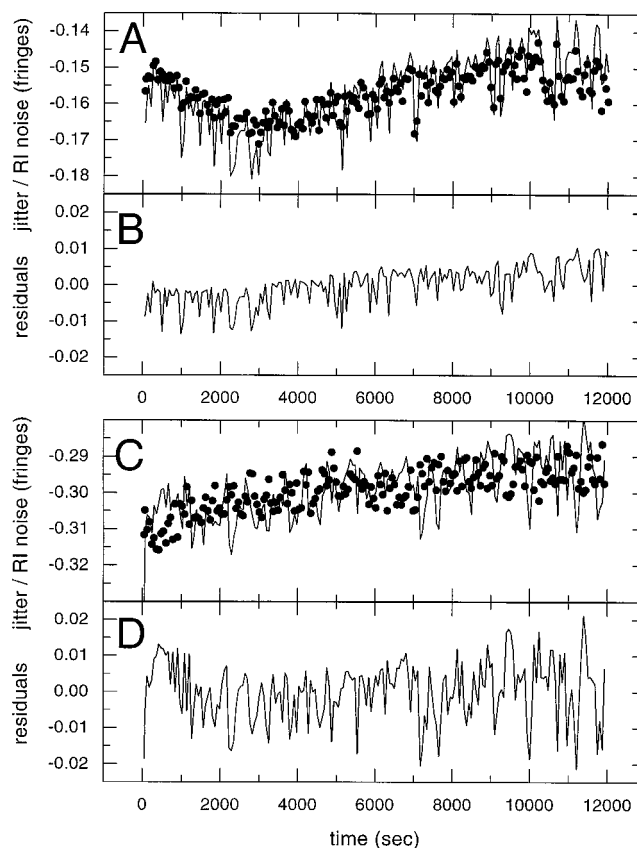

FIGURE 3   Comparison of the calculated radial-invariant noise component (*lines*) with the calculated jitter from least-squares air-to-air alignment (*symbols*), after elimination of a constant offset. (*A*) Results from the complete analysis of the myoglobin experiment, a subset of which is shown in Fig. 1. (*B*) Difference of the results from both methods, with rms deviation of 0.0052 fringes. (*C*) Results from the complete analysis an experiment with gamma globulin, a subset of which is shown in Fig. 4. (*D*) Difference of the results from both methods, with a rms deviation of 0.0084.

fraction of 19% refractive index contribution from species 2. These values seem consistent with the g(s*) profiles, and lead to a rmsd of 0.0072 fringes (which is significantly lower than the rmsd of 0.0086 fringes obtained from a single species model). This demonstrates consistency of the Lamm analysis with the g(s*) results. These results are confirmed by the analysis of the corresponding absorbance profiles (data not shown). Air-to-air alignment of the full data set of 200 scans before analysis, in place of the algebraic RI noise calculation, significantly increased the rms error to 0.0108, which is even higher than the value of 0.0099 fringes obtained without any offset correction. This finding correlates with the relative large deviation in the high-frequency part of the air-to-air and the calculated RI noise corrections (Fig. 3 *D*).

Finally, we analyzed the sedimentation profiles of a mixture of myoglobin and gamma globulin as detected with the interference optics (Fig. 5). Without further fitting for the sedimentation coefficients and using the results obtained before (see above), unraveling the systematic noise compo-
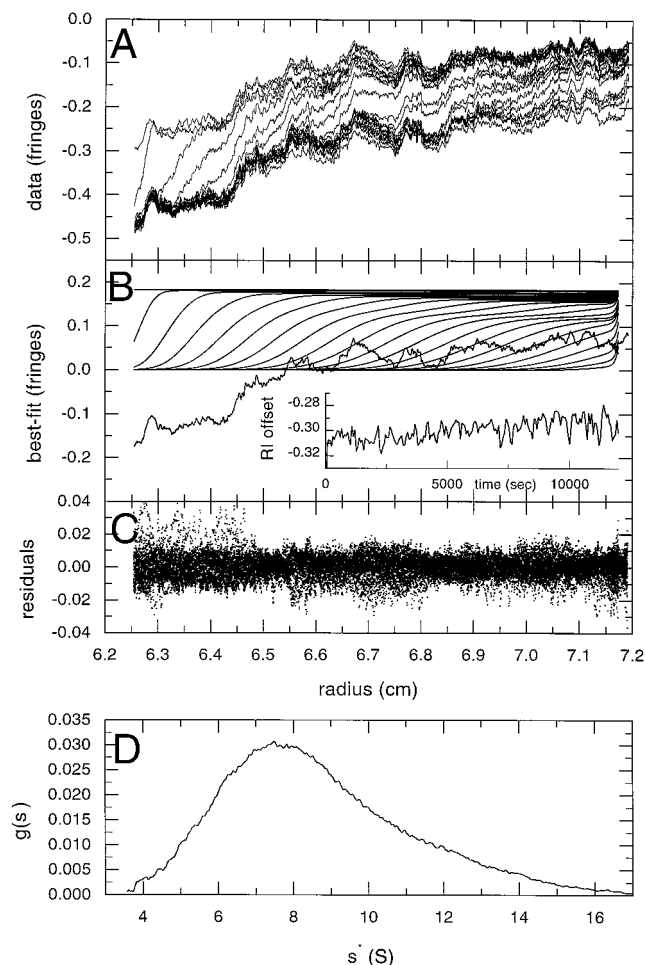
FIGURE 4 Analysis of the interference data of the sedimentation gamma globulin. (*A*) Experimental fringe displacement profiles. For clarity, only every 10th scan of the total of 200 is shown. (*B*) Best-fit decomposition into the calculated sedimentation profiles of two noninteracting components (constraining the molar masses to $M_1 = 160,000$ Da and $M_2 = 320,000$ Da, respectively, and the partial specific volumes to 0.734 ml/g). The inset shows noise contributions as a function of time. The calculated TI noise is shown below the calculated solutions of the Lamm equation. Best-fit sedimentation coefficients from the analysis of all 200 scans are $s_{\gamma 1} = 7.72$ S and $s_{\gamma 2} = 11.6$ S, with a loading concentration of 0.178 fringes and a 19% relative contribution from species 2 (the best single species fit leads to $s_{w\gamma} = 7.90$ S at a rmsd of 0.0086). (*C*) Residuals of the fit, with a rmsd of 0.0072 fringes in the complete analysis of the 200 scans. (*D*) g(s*) transformation of the scans 40 to 70. The resulting weight average sedimentation coefficient is 8.64 ± 0.19 S, with a loading concentration of 0.17 fringes.

FIGURE 5 (*A*) Experimental sedimentation profiles of a mixture of myoglobin and gamma globulin. Only every 10th scan of the total of 200 is shown. (*B*) Decomposition into the calculated sedimentation profiles of myoglobin, gamma globulin, and a dimeric aggregate of gamma globulin, with sedimentation coefficients predetermined as shown in Figs. 1 and 3, and the systematic noise components. The inset shows the RI noise contributions as a function of time. The calculated TI noise is shown below the calculated solutions of the Lamm equation. (*C*) Residuals of the decomposition, which leads to a rmsd of 0.0051 fringes in the complete analysis of the 200 scans.

nents led to a rmsd of 0.0051 fringes. This excellent description of the data further demonstrates the validity of this treatment of the systematic noise components of interference data.

The calculation of systematic TI noise components can be applied also to the analysis of sedimentation equilibrium. If scans are taken during the approach to equilibrium for the identification of TI noise components, the sedimentation equilibrium profiles can be corrected for the contribution of the TI noise. This can be particularly useful for eliminating
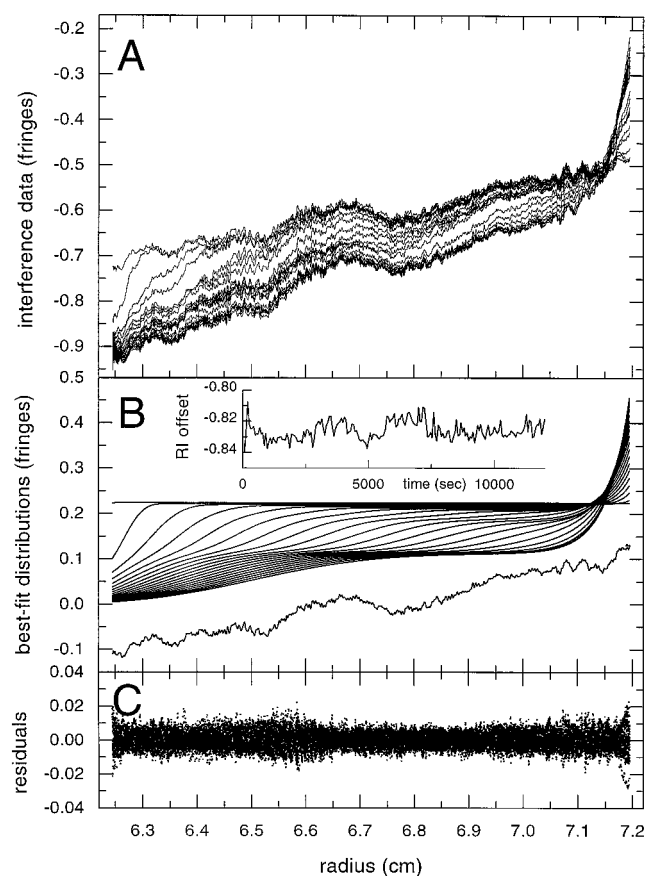
the baseline noise from interference equilibrium data, as well as eliminating the effects of imperfections of the window from absorbance equilibrium profiles. Fig. 6 shows the approach to equilibrium for a protein in a monomer-dimer equilibrium. Visual inspection of the approach to equilibrium data and of the equilibrium profiles clearly shows a distinct time-invariant background profile. Modeling the approach to equilibrium data was performed with a superposition of Lamm equation solutions, with the strategy of using a very large model by allowing for several components with floating sedimentation and diffusion coefficient, as well as a floating time offset for start of the simulated sedimentation. Ideally, the model should allow for solutions of the Lamm equation that approximately span the whole function space of Lamm equation solutions. While this analysis is necessarily ill-conditioned in terms of the obtained *s*- and *D*-values that will not be interpreted, this property is utilized to avoid the introduction of systematic
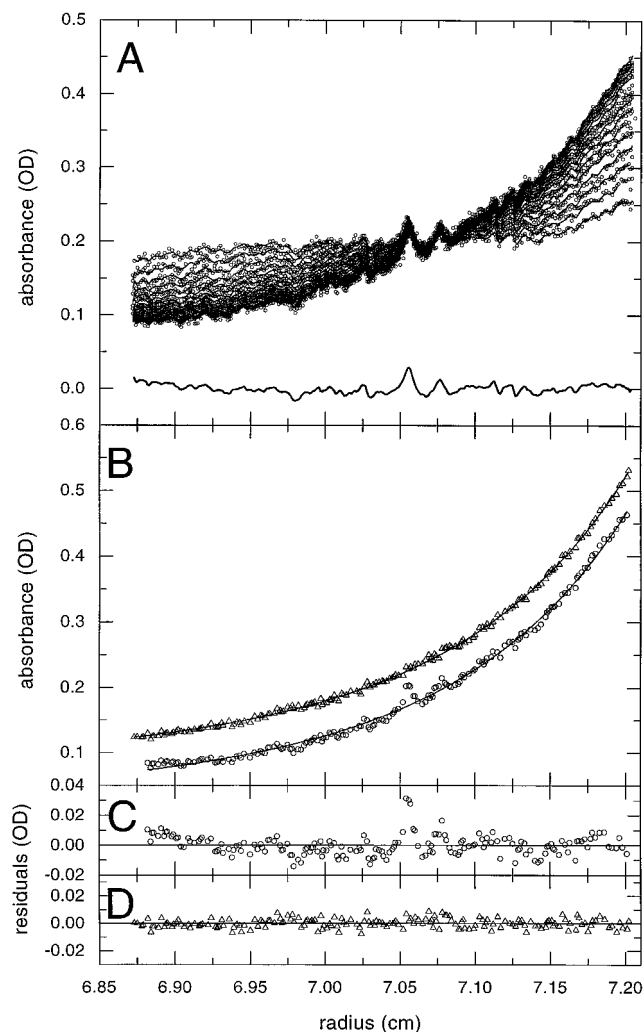
FIGURE 6 (*A*) Absorbance profiles at 230 nm of the approach to equilibrium of a protein undergoing a reversible monomer-dimer self-association (*circles*). Best-fit Lamm equation model (*solid lines*) taking TI noise contributions into account (solid line below with average zero). (*B*) Raw sedimentation equilibrium data (*circles*), equilibrium data corrected for the TI noise contribution shown above (*triangles*), and best-fit global sedimentation equilibrium analysis of both data sets (*solid lines*). The corrected sedimentation equilibrium profile is offset by 0.05 OD for clarity. The residuals of these fits are shown in (*C*) and (*D*), respectively.

error in the calculated TI background by using a wrong model. Fitting with an ill-conditioned and very large Lamm equation model effectively resembles a constrained smoothing operation that allows the well-conditioned determination of the higher frequency parts of the systematic TI noise, which is shown in Fig. 6 *A*. It is noteworthy that the baseline contribution exhibits features similar to the residuals obtained from the analysis of the uncorrected sedimentation equilibrium data (Fig. 6 *C*). The analysis of the difference of the sedimentation equilibrium profiles and the calculated TI noise leads to residuals that are significantly smaller and appear more random (Fig. 6 *D*). Importantly, as a consequence of the improved quality of the data, the error estimate for the dimerization constant obtained was reduced by a factor of 2.

DISCUSSION

We have presented an algebraic method for the consideration of systematic time-invariant and radial-invariant noise components in the least-squares analysis of analytical ultracentrifuge data from sedimentation velocity experiments. The origin of some of these systematic noise components are probably small thermal or mechanical instabilities, since changes in the optical path length by only a few nanometers can already lead to a significant fringe shift. This could include thermal gradients in the heat sink and across the condensing lens, as well as slow changes in the geometry of the optical configuration, such as motions of the recording plate, the lens, or the light source (T. M. Laue, personal communication).

The described approach is very efficient and yields consistently good results, as demonstrated by the exemplary analyses of the sedimentation profiles of myoglobin (Fig. 1) and gamma globulin (Fig. 4), which led to rms deviations of only 0.0065 and 0.0072 fringes, respectively. That the calculated noise components are realistic is indicated by the analysis of the sedimentation profiles of a mixture of the two proteins (Fig. 5). Decomposition of the data directly into the Lamm equation solutions with predetermined coefficients and into the systematic noise components leads to a rms deviation of only 0.0051 fringes. These results also demonstrate that the presented analytical approach can fully exploit the high sensitivity that can be experimentally achieved with the interference optical system. Likewise, that the calculated sedimentation coefficients are reliable is demonstrated by the excellent agreement obtained for results from experiments performed on either optical system (Fig. 2), as well as their consistency with the results from *dc/dt* analysis (Fig. 4 *D*).

Direct calculation of TI and RI noise through separation of linear and nonlinear parameters is very versatile. Although the examples shown were based on direct fitting of numerical solutions of the Lamm equation, no reference is made in the derivation to the nature of the functions $L_{n,i}^{(k)}$, and they can represent approximate analytical solutions of the Lamm equation, or even sets of dependent sedimentation profiles of interacting systems. Therefore, the presented method can be readily incorporated into any other method for direct boundary fitting and applied to the analysis of interference optical ultracentrifuge data. Obviously, the least-squares calculation of RI noise components can incorporate, without any further difficulty, integral fringe shifts that occasionally were found in a subset of the scans (and removed for presentation purposes). Furthermore, while the consideration of time-invariant noise components is crucial for interference data, it can also be advantageous in the analysis of absorbance data, where imperfections in the windows can also lead to small time-invariant signal offsets, as indicated in Fig. 2. It can be successfully employed for the identification of the time-invariant background from the approach to equilibrium data, which subsequently can be used for the correction of sedimentation equilibrium pro-

files. Here, advantage is taken of the fact that low-speed data have less specific hydrodynamic information (Schuck, 1998); this allows an analysis which is ill-conditioned with respect to the Lamm equation parameters, but is well-conditioned with respect to the high-frequency TI noise. While this technique could be particularly useful in the analysis of interference optical sedimentation equilibrium experiments, its effectiveness has been demonstrated for absorbance optical data in Fig. 6.

While fitting for the systematic noise components introduces a very large number of unknowns, the results of the error analyses described above indicate that they correlate only weakly with the parameters governing the macromolecular sedimentation, and therefore only slightly decrease the precision of the s-value. In part, this appears to be due to the intrinsic smoothness of the macromolecular sedimentation profiles, which cannot describe and correlate with the high spatial frequency of the TI noise. However, in particular for sedimentation profiles of small molecules that do not exhibit distinct sedimentation boundaries, and in the analysis of small datasets, some slight correlation can exist between the sedimentation parameters and the low spatial frequency part of the calculated TI noise. It should be considered, though, that all acceptable models identified with the presented method for the direct calculation of the TI noise components obviously also provide an acceptable description of the evolution of the time-difference of the scans, which is the basis of current alternative methods of interference data analysis. As a consequence, the possible slight correlation of sedimentation parameters with low spatial frequency parts of the TI noise is a problem that is inherently introduced by local offsets in the experimental data acquisition, and not by the strategy of the data analysis.

In the analysis of interference optical ultracentrifuge data, other, more empirical methods for the elimination of TI and RI measurement errors have been used. This includes the subtraction of a separately measured background scan for correcting TI noise contributions to the data. However, while a simple subtraction would amplify the random noise component, it can also be insufficient if depositions of oil from the drive or moisture accumulate on the windows and change the "time-invariant" background, as pointed out by Stafford (1994). One method employed for the elimination of radial-invariant offset is the alignment of the scans in the air-to-air region above the solution column. When we compared the results of this method with the calculated RI noise components (Fig. 3), we found that both methods correspond well in the description of slow drifts of this signal offset. However, we found relatively large discrepancies in the fluctuations of the offset on a short time scale (in the order of minutes). These high-frequency contributions to the RI noise are of the same magnitude as the slow drifts, and are statistically significant. Given the low random noise of the data (~0.005–0.007 fringes) and considering that the offsets were calculated by air-to-air alignment on the basis of 100 data points per scan, or even >1000 data points per scan in the case of the algebraic calculation, respectively,

the statistical error of the calculated offsets is considerably smaller than the observed magnitude of 0.005–0.01 fringes of its high-frequency part. A possible explanation of this error could be the presence of small mechanical vibrations in the optical configuration. Air-to-air alignment was not successful as a means of correction for this measurement error. (Also, slight flexing of some isolated scans was observed in the residuals plot, which could result in limited effectiveness of the air-to-air alignment.) Unfortunately, the need for consideration of RI noise in the data analysis leads to a slight decrease of the statistical accuracy of the derived s-values.

Compared with the time-derivative (dc/dt) analysis, used in the g(s*) transformation, the method for direct analysis presented in this paper has several advantages: first, it does not require the differentiation of discrete and noisy data and avoids the related well-known problems of error amplification. Therefore, the direct analysis is statistically advantageous. Second, it allows direct comparison of the model with the raw data, and therefore inspection of the residuals and the application of statistical tests of goodness-of-fit. Although, according to Eq. 2, systematic "collective" deviations will not show in the residuals but are transferred into the calculated TI noise, systematic deviations of the model from the experimental data in the individual scans can still be observed and taken as a criterion for goodness-of-fit. Third, it allows the calculation of RI noise components, as opposed to their extrapolation from the air-to-air alignment. The air-to-air alignment was found to describe well the slow drifts of the offset, but not the higher-frequency components, which we found to be of the same magnitude and statistically highly significant (see above). Finally, the direct analysis is more general and can be easily combined with any Lamm equation solution, including for those small molecules or for interacting systems.

An alternative strategy for the application of Lamm equation fitting for small molecules or interacting systems to interference optical data is a Δc-approach (Stafford, 1998; J. S. Philo, personal communication). This is a modification of the time-derivative method by fitting of finite time differences of the Lamm equation solutions to the respective time differences of the data. While the substitution of derivatives with differences diminishes some of the problems connected with error amplification in the calculation of numerical derivatives, it inherits the drawbacks of the dc/dt methods with respect to the inspection and analysis of the goodness-of-fit, and with respect to the RI noise that is time-dependent.

One more fundamental difference between the g(s*) method and the Lamm equation approach is that the g(s*) transformation is relatively model-independent. In this property, it is related to the method of van Holde and Weischet (Demeler et al., 1997; van Holde and Weischet, 1978), which allows the performance of rapid diagnostics of the sample. These model-free approaches can be highly useful for the identification of the number of species, associative behavior, concentration dependence, and estimation

of sedimentation coefficients, information which we believe can be very valuable for the subsequent incorporation into a detailed Lamm equation model. While the *dc/dt* approach of the g(s*) transformation is intrinsically insensitive to TI noise, the application of the van Holde-Weischet analysis to interference optical data would require model-free methods for the elimination of TI and RI noise. Once a model has been identified, however, the treatment of TI and RI noise, as presented here, in combination with the recently described methods for Lamm equation analysis allows for the direct comparison of experimental and fitted data, and enables the full exploitation of the high precision of interference optical ultracentrifuge data.

REFERENCES

Beechem, J. M. 1992. Global analysis of biochemical and biophysical data. *Methods Enzymol.* 210:37–54.

Behlke, J., and O. Ristau. 1997. Molecular mass determination by sedimentation velocity experiments and direct fitting of the concentration profiles. *Biophys. J.* 72:428–434.

Bethune, J. L., and G. Kegeles. 1961a. Countercurrent distribution of chemically reacting systems. I. Polymerization. *J. Phys. Chem.* 65: 433–438.

Bethune, J. L., and G. Kegeles. 1961b. Countercurrent distribution of chemically reacting systems. III. Analogs of moving boundary electrophoresis and sedimentation. *J. Phys. Chem.* 65:1761–1764.

Bevington, P. R., and D. K. Robinson. 1992. Data Reduction and Error Analysis for the Physical Sciences. Mc-Graw-Hill, New York.

Bloomfield, V., W. O. Dalton, and K. E. Van Holde. 1967. Frictional coefficients of multisubunit structures. *Biopolymers.* 5:135–148.

Byron, O. 1997. Construction of hydrodynamic bead models from high-resolution x-ray crystallographic or nuclear magnetic resonance data. *Biophys. J.* 72:408–415.

Cann, J. R., and G. Kegeles. 1974. Theory of sedimentation for kinetically controlled dimerization reactions. *Biochemistry.* 13:1868–1874.

Claverie, J.-M. 1976. Sedimentation of generalized systems of interacting particles. III. Concentration-dependent sedimentation and extension to other transport methods. *Biopolymers.* 15:843–857.

Correia, J. J. 1998. Sedimentation velocity data analysis methods: what, when and why? *Chemtracts: Biochemistry and Molecular Biology.* in press.

Cox, D. J. 1969. Computer simulation of sedimentation in the ultracentrifuge. IV. Velocity sedimentation of self-associating solutes. *Arch. Biochem. Biophys.* 129:106–123.

Cox, D. J., and R. S. Dale. 1981. Simulation of transport experiments for interacting systems. *In* Protein-Protein Interactions. C. Frieden and L. W. Nichol, editors. Wiley, New York.

de la Torre, J. G. 1992. Sedimentation coefficients of complex biological particles. *In* Analytical Ultracentrifugation in Biochemistry and Polymer Science. S. E. Harding, A. J. Rowe, and J. C. Horton, editors. Royal Society of Chemistry, Cambridge, UK. 333–358.

Demeler, B., and H. Saber. 1998. Determination of molecular parameters by fitting sedimentation data to finite element solutions of the Lamm equation. *Biophys. J.* 74:444–454.

Demeler, B., H. Saber, and J. C. Hansen. 1997. Identification and interpretation of complexity in sedimentation velocity boundaries. *Biophys. J.* 72:397–407.

Fujita, H. 1962. Mathematical Theory of Sedimentation Analysis. Academic Press, New York.

Gilbert, L. M., and G. A. Gilbert. 1973. Sedimentation velocity measurement of protein association. *Methods Enzymol.* 27:273–296.

Goad, W. B., and J. R. Cann. 1969. Theory of sedimentation of interacting systems. *Ann. N.Y. Acad. Sci.* 164:172–182.

Holladay, L. A. 1979. An approximate solution to the Lamm equation. *Biophys. Chem.* 10:187–190.

Johnson, M. L., and M. Straume. 1994. Comments on the analysis of sedimentation equilibrium experiments. *In* Modern Analytical Ultracentrifugation. T. M. Schuster and T. M. Laue, editors. Birkhäuser, Boston. 37–65.

Lamm, O. 1929. Die Differentialgleichung der Ultrazentrifugierung. *Ark. Mat. Astr. Fys.* 21B(2):1–4.

Laue, T. M. 1994. An on-line interferometer for the XL-A ultracentrifuge. *Prog. Coll. Polym. Sci.* 94:74–81.

Laue, T. M. 1997. Advances in sedimentation velocity analysis. *Biophys. J.* 72:395–396.

Laue, T. M., R. A. Domanik, and D. A. Yphantis. 1983. Rapid precision interferometry for the analytical ultracentrifuge. I. A laser controller based on a phase-lock-loop circuit. *Anal. Biochem.* 131:220–231.

Laue, T. M., D. A. Yphantis, and D. G. Rhodes. 1984. Rapid precision interferometry for the analytical ultracentrifuge. III. Determination of period of rotation, frequency of rotation, and elapsed time. *Anal. Biochem.* 143:103–112.

Lawson, C. L., and R. J. Hanson. 1974. Solving Least Squares Problems. Prentice-Hall, Englewood Cliffs, NJ.

Philo, J. S. 1997. An improved function for fitting sedimentation velocity data for low molecular weight solutes. *Biophys. J.* 72:435–444.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. Numerical Recipes in C. University Press, Cambridge.

Ruhe, A., and P. Å. Wedin. 1980. Algorithms for separable nonlinear least squares problems. *SIAM Rev.* 22:318–337.

Schuck, P. 1998. Sedimentation analysis of noninteracting and self-associating solutes using numerical solutions to the Lamm equation. *Biophys. J.* 75:1503–1512.

Schuck, P., C. E. MacPhee, and G. J. Howlett. 1998. Determination of sedimentation coefficients for small proteins. *Biophys. J.* 74:466–474.

Schuster, T. M., and J. M. Toedt. 1996. New revolutions in the evolution of analytical ultracentrifugation. *Curr. Opin. Struct. Biol.* 6:650–658.

Stafford, W. F. 1992. Boundary analysis in sedimentation transport experiments: a procedure for obtaining sedimentation coefficient distributions using the time derivative of the concentration profile. *Anal. Biochem.* 203:1–7.

Stafford, W. F. 1994. Boundary analysis in sedimentation velocity experiments. *Methods Enzymol.* 240:478–501.

Stafford, W. F. 1997. Sedimentation velocity spins a new weave for an old fabric. *Curr. Opin. Biotechnol.* 8:14–24.

Stafford, W. F. 1998. Time difference sedimentation velocity analysis of rapidly reversible interacting systems: determination of equilibrium constants by global nonlinear curve fitting procedures. *Biophys. J.* 74:301*a*. (Abstr.).

Svedberg, T., and K. O. Pedersen. 1940. The Ultracentrifuge. Oxford University Press, London.

van Holde, K. E., and W. O. Weischet. 1978. Boundary analysis of sedimentation velocity experiments with monodisperse and paucidisperse solutes. *Biopolymers.* 17:1387–1403.

Yphantis, D. A. 1984. Sedimentation coefficient distributions from ultracentrifuge interferograms. *Biophys. J.* 45:324*a*. (Abstr.).

Yphantis, D. A., J. W. Lary, W. F. Stafford, S. Liu, P. H. Olsen, D. B. Hayes, T. P. Moody, T. M. Ridgeway, D. A. Lyons, and T. M. Laue. 1994. On-line data acquisition for the Raleigh interference optical system of the analytical ultracentrifuge. *In* Modern Analytical Ultracentrifugation. T. M. Schuster and T. M. Laue, editors. Birkhäuser, Boston. 209–226.

Yphantis, D. A., T. M. Laue, and I. Anderson. 1984. Rapid precision interferometry for the analytical ultracentrifuge. II. A laser controller based on a rate-multiplying circuit. *Anal. Biochem.* 143:95–102.