shapes of the curves in the figure are not changed but the absolute values of the χ parameter are decreased, which decrease the values of the β parameter. However the conclusion of the similarity of results calculated from the SPT and CST remains unaffected.

Conclusions

Remarkably similar predictions are obtained using the SPT and the CST plus Flory model of liquid state. The CST has drawn attention to the important role played by free volume in polymer solution thermodynamics. It seems that the SPT is also successful in treating free volume effects. We believe that the great popularity of the solubility parameter approach, particularly in industrial laboratories, is entirely justified.

Acknowledgment. The support of the National Research Council of Canada is gratefully acknowledged.

Analysis of Polydisperse Systems at Sedimentation Equilibrium. I. Simple Solvent Systems

Arthur Rosenthal

Chemistry Department, State University of New York, Albany, New York 12208.
Received July 8, 1970

ABSTRACT: An analytical procedure is developed which permits the complete characterization of polydisperse systems. The computational algorithm involves a nonlinear "least squares" determination of the unknown coefficients for an equation describing sedimentation equilibrium. It is demonstrated that by the appropriate modification of the functional form or the selective application of constraints on the coefficients the same basic algorithm can encompass the requirements of associating heterogeneous systems with sufficient flexibility to accept the data forms produced by the interference, photoelectric scanning, and schlieren optical systems. The possibility of treating several other aspects of the analysis of systems at sedimentation equilibrium is noted as well as the applicability of the basic computational procedure to general problems of data analysis.

The analytical ultracentrifuge is an experimental tool of great power. Not the least aspect of this instrument is its potential for determining molecular weights of macromolecules. In the absence of significant heterogeneity, molecular weights can be determined with ease and reliability. Such systems are, however, neither extremely common nor are they particularly exciting. A large proportion of real systems will contain contaminants which may or may not be experimentally important. Even more significant are systems which involve association as a function of concentration and/or solvent composition. Two approaches have been available for the interpretation of sedimentation equilibrium experiments in these cases. The most generally employed analytical procedures have involved the determination of apparent average molecular weight distributions.[1-9] For the special cases where association occurs these distributions have been analyzed for the molecular weights and association constants. Likewise for systems involving self-association the direct determination of the monomer molecular weight and association constants by fitting the data to specific model systems has been proposed.[10,11]

It is the purpose of this paper to demonstrate that an analysis of the total observed radial distribution can provide a direct determination of the molecular weights and concentration distributions of each component not only for associating systems but also for the general case of a heterogeneous system.[12] At the same time the considerations required for optimum experimental design will be considered. For a thorough analysis of sedimentation phenomena it would be necessary to include the possibility of preferential interactions and of density gradients not ordinarily allowed for in simple analyses. In the present work we will only indicate the direction to be taken for these two considerations as they will be considered in depth subsequently.

Theory

The sedimentation equilibrium relationship, presented by Svedberg and Pederson[14] in their classic treatise on the ultracentrifuge, has been the basis for most equilibrium analyses performed. This relationship is normally expressed according to eq 1, in which M is the

(1) D. A. Yphantis, *Biochemistry*, **3**, 297 (1964).
(2) E. T. Adams, Jr., and J. W. Williams, *J. Amer. Chem. Soc.*, **86**, 3454 (1964).
(3) E. T. Adams, Jr., *Biochemistry*, **4**, 1646 (1965).
(4) E. T. Adams, Jr., *ibid.*, **6**, 1864 (1967).
(5) L. W. Nichol and A. G. Ogston, *J. Phys. Chem.*, **69**, 4365 (1965).
(6) H. W. Osterhoudt and J. W. Williams, *ibid.*, **69**, 1050 (1965).
(7) E. G. Richards, D. C. Teller, and H. K. Schachman, *Biochemistry*, **7**, 1054 (1968).
(8) Th. G. Scholte, *Ann. N. Y. Acad. Sci.*, **164**, 156 (1969).
(9) A. Rosenthal, Doctoral Dissertation, University of California at Berkeley, 1969.
(10) E. T. Adams, Jr., *Ann. N. Y. Acad. Sci.*, **164**, 226 (1969).
(11) (a) D. E. Roark and D. A. Yphantis, *ibid.*, **164**, 245 (1969); (b) K. E. VanHolde, G. P. Rossetti, and R. D. Dyson, *ibid.*, **164**, 279 (1969).
(12) Haschemeyer and Bowers[13] have published a similar treatment since this work was completed but imply that the analysis is more limited than need be.
(13) R. H. Haschemeyer and W. F. Bowers, *Biochemistry*, **9**, 435 (1970).
(14) T. Svedberg and K. O. Pederson, "The Ultracentrifuge," Oxford University Press, New York, N. Y., 1940.

$$M = \frac{2RT\mathrm{d}\ln c}{(1 - \bar{v}\rho)\omega^2\mathrm{d}r^2} \quad (1)$$

molecular weight of the macromolecule, R is the gas constant in ergs/mol deg, T is the temperature in °K, c is the concentration of the macromolecule at the radial locus r in arbitrary units, $\bar{v}$ is the partial volume of the macromolecule in cc/g, ρ is the density of the solution in g/cc, ω is the angular velocity in radians/sec, and r is the distance from the center of the rotor in cm. There are several possible complications to eq 1: preferential interactions, redistribution of cosolute (variable ρ), nonideality, and heterogeneity. Of these we will develop only the last in this first paper.

Analysis of Heterogeneity

Equation 1 is written specifically for systems involving only one, nonaggregating, macromolecule which does not interact with the solvent. In the event of more than one component, eq 1 applies to each component as

$$M_i = \frac{2RT\mathrm{d}\ln c_i}{(1 - \bar{v}_i\rho)\omega^2\mathrm{d}r^2} \quad (1a)$$

where the subscript i indicated the ith macromolecular component. On rearrangement of eq 1a one obtains

$$M_i c_i = \frac{2RT\mathrm{d}c_i}{(1 - \bar{v}_i\rho)\omega^2\mathrm{d}r^2} \quad (2)$$

If $\bar{v}_i$ is assumed to be the same for all i, eq 2 can be summed over all i to produce

$$\sum_i M_i c_i = \frac{2RT\sum_i \mathrm{d}c_i}{(1 - \bar{v}\rho)\omega^2\mathrm{d}r^2} \quad (3)$$

Since the summation of $\mathrm{d}c_i$ is simply the total $\mathrm{d}c$ and the weight average molecular weight is defined, according to eq 4, as

$$M_w = \frac{\sum_i M_i c_i}{\sum_i c_i} \quad (4)$$

in which the summation of c_i is simply c total, one can write

$$M_{w,r} = \frac{2RT\mathrm{d}\ln c}{(1 - \bar{v}\rho)\omega^2\mathrm{d}r^2} \quad (5)$$

Equation 5 is most attractive. One need only plot the natural logarithm of the total concentration against the square of the radius and the weight average molecular weight at any radial locus may be obtained directly from the slope at that radial locus. For all of the appeal and ease of this procedure it has several distinct disadvantages. An obvious and usually minor problem is that it must be assumed that all $\bar{v}_i$ and $(\partial n/\partial c_i)_{T,P,c_i}$,[15] where n represents the refractive index, are the same. A second difficulty lies in the requirement for knowing the absolute concentration in order to evaluate the logarithm. The interferometric optical system provides the most precise means of determining concentration in the ultracentrifuge.[9] It, sadly enough, only indicates increments of concentration. While several procedures, both experimental and computational, have been put forth to circumvent this problem,[1,7,9] they all leave something to be desired. The experimental systems involve either additional experimental effort or a substantial restriction on the range of experimental conditions which can be used. The computational approach, while reasonably reliable in general, can break down on occasion and requires an approximate foreknowledge of the system composition. Another complication, in the long run quite possibly the most serious, is more subtle. The slope required for eq 5 must be determined by some type of approximation, generally a least-squares fitting of the curve to obtain coefficients, followed by differentiation to obtain the derivative[7] needed to evaluate M_w. Alternatively, orthogonal least-squares fitting may be done which produces a much better derivative, but is still a power series approximation.[9] There is nothing wrong with this unless one is seeking the molecular weights of the various species present, instead of just the weight average distribution. As a result of forcing the ln c *vs.* r^2 relationship into a least-squares power series one has also imposed a unique bias onto the apparent weight average molecular weight evaluation which prohibits a truly accurate secondary analysis of the weight average molecular weight as a function of the concentration as would be required for a determination of the individual components.

It is possible to perform the analysis initially in such a way as to obviate the foregoing difficulties. Integration of eq 1 produces

$$c_r = y_r - \delta = c_{ref}e^{A M(1-\bar{v}\rho)\Delta r^2} \quad (6)$$

in which $y_r - \delta$ is the concentration, c_r, at the radial locus r, y_r being the observed increment and δ the necessary correction to adjust to true c_r, c_{ref} is the concentration at an arbitrary reference locus, Δr^2 is described by

$$\Delta r^2 = r^2 - r_{ref}^2 \quad (6a)$$

and the parameter A is defined as

$$A \equiv \frac{\omega^2}{2RT} \quad (6b)$$

For the case of heterogeneity eq 6 can be expressed as

$$c_r = \sum_i c_{i,r} = y_r - \delta = \sum_i c_{i,ref}e^{A M_i(1-\bar{v}_i\rho)\Delta r^2} \quad (7)$$

in which the subscript i again represents the ith component. For convenience r_{ref} may be chosen as the meniscus, r_m, in which case $c_{i,ref}$ becomes $c_{i,m}$ and Δr^2 becomes $r^2 - r_m^2$. Equation 7 does not impose a requirement for true c or for equal values of the $\bar{v}_i$ or $(\partial n/\partial c_i)_{T,P,c}$[16] terms since these are simply additional parameters. The difficulty of applying eq 7 is that analytical solutions of nonlinear equations are not, in general, particularly facile.

It is possible, however, to develop an iterative least-

(15) For the interferometric and schlieren optical systems the assumption is made that c is linearly related to the refractive increment. With the absorption optical system the corresponding assumption is that the specific absorbance is identical for all protein species.

(16) This term does not enter explicitly but is implicit in the evaluation of c_i.

squares analysis[17] which will generally converge on a good estimate with satisfactory rapidity. Equation 7 can be expressed as

$$y_r = f(K_j) \tag{7a}$$

in which the set of K_j represents the j unknown coefficients, δ, $c_{i,m}$, and M_i (for the present the $\bar{v}_i$ are being assumed equal).[18] One can then enter initial approximations, denoted as prime for the unknown coefficients and perform a single term Taylor expansion for f as

$$y_r - d_r = f_r(K_j') + \sum_j \Delta K_j \left(\frac{\partial f_r}{\partial K_j'}\right) \tag{8}$$

in which ΔK_j stands for the correction required to achieve the next order of approximation, for example

$$\begin{aligned} \delta'' &= \delta' + \Delta\delta \\ c_{i,m}'' &= c_{i,m}' + \Delta c_{i,m} \\ M_i'' &= M_i' + \Delta M_i \end{aligned} \tag{9}$$

and d_r is the deviation of f_r from y_r. The definition of a least-squares fit is that the summation of d_r^2 over all r becomes a minimum. This can be written, in the present notation, as

$$\sum_r d_r^2 = \sum_r \left[\sum_j \Delta K_j \left(\frac{\partial f_r}{\partial K_j'}\right) - (y_r - f_r(K_j'))\right]^2 = \text{minimum} \tag{10}$$

The set of derivatives of Σd_r^2 with respect to the several ΔK_j terms can then be stated as

$$\frac{\partial(\Sigma d_r^2)}{\partial \Delta K_j} = 2\sum_r \left[\sum_j \Delta K_j \left(\frac{\partial f_r}{\partial K_j'}\right) - (y_r - f_r(K_j'))\right]\left(\frac{\partial f_r}{\partial K_j'}\right) = 0 \tag{11}$$

where $j = 1, 2, \ldots n$. Since there are n unknowns in the set of eq 11 and n equations, the determination of the next order corrections for each of the coefficients of the original nonlinear equation involves a simple matrix inversion.

This solution is not only relatively easy to use, it is also extremely powerful in that it allows great facility in the application of constraints or in altering the functional forms. The coefficients entered into the matrix routine for correction are not defined arbitrarily, as with a power series, but rather are defined explicitly by the statement of the function. As a consequence, one may impose constraints of almost any form with ease. For example, a coefficient, c, may be forced to take only positive values by defining it as the square of a different coefficient, b, or defined to a fixed value by assignment as a determined constant, etc.

(17) H. Margenau and G. M. Murphy, "The Mathematics of Physics and Chemistry," 2nd ed, Van Nostrand, Princeton, N. J., 1956, pp 516–519.

(18) This was done for ease of data simulation. It was not a requirement of the algorithm. Since concentrations have all been expressed in fringe values, *i.e.*, refractive increments, there is no assumption at this stage of the analysis, regarding $(\partial n/\partial c_i)_{P,T,c_j}$. This assumption is only required when one seeks to relate the concentration of one component to that of another under circumstances where both must be on the same concentration scale. This would be the case in the assessment of association constants, for example.

A minor modification of eq 7 will illustrate how the function itself can be used to apply a constraint. For self-associating systems one can write

$$M_i = iM_1 \tag{12}$$

which, with an assumption of uniform $\bar{v}_i$, permits one to reexpress eq 7 as

$$c_r = y_r - \delta = \sum_i c_{i,m} e^{iAM_1(1-\bar{v}\rho)\Delta r^2} \tag{13}$$

Not only does eq 13 have far fewer variables than eq 7, but minor modifications permit a rearrangement to a linear form.[9] This is not as advantageous as it might seem since in its nonlinear form it is possible to select the values of i to be considered, providing, of course, that there exists some basis for the selection.

Computational Considerations

Statistics. When attempting to translate any set of data into a functional representation, even a representation which is based on a theoretically correct functional form, there are several problems which must be considered. One must assess the potential significance of the individual points, the reasonable level to which the data can be "fit" with any statistical validity, the experimental conditions which will produce optimal results, the reliability of the analytical procedure involved on known systems, and finally the extent to which a solution, once obtained, can be regarded as a meaningful or unique solution.

The relative contributions of individual points is clearly not a constant. It can be demonstrated that the error in the molecular weights depends on $(\Delta r^2 e^{A\Delta r^2})^{-1}$ times the error in y_r. Since the absolute magnitude of the probable error in y_r is independent of y_r, it is apparent that $\Delta r^2 e^{A\Delta r^2}$ provides a good weighting factor for each data point.

In the construction of the analytical arguments one premise was the minimization of the root mean square (rms) error. The true minimum value in this parameter may or may not be fully significant within the context of the experimental system. Indeed, there is no justification in reducing the rms error below the probable error of the readings. As a result this error level is used as a termination level for the analysis. While it is in principle possible to develop statistical arguments for the other questions posed, it is, in practice, more direct to consider them empirically; therefore, discussion of these items will be deferred until the concluding sections.

Matrix Analysis. Equation 11 can be represented in matrix form as

$$\begin{bmatrix} fK_1 \cdot fK_1 & fK_2 \cdot fK_1 & \cdots & fK_n \cdot fK_1 \\ fK_1 \cdot fK_2 & fK_2 \cdot fK_2 & \cdots & fK_n \cdot fK_2 \\ \cdots & \cdots & \cdots & \cdots \\ fK_1 \cdot fK_n & fK_2 \cdot fK_n & \cdots & fK_n \cdot fK_n \end{bmatrix} \begin{bmatrix} \Delta K_1 \\ \Delta K_2 \\ \cdots \\ \Delta K_n \end{bmatrix} = \begin{bmatrix} D \cdot fK_1 \\ D \cdot fK_2 \\ \cdots \\ D \cdot fK_n \end{bmatrix} \tag{14}$$

where

$$fK_i = \sum_r \left(\frac{\partial f_r}{\partial K_i}\right) \tag{14a}$$

and

$$D \cdot fK_i = \sum_r (y_r - f_r)\left(\frac{\partial f_r}{\partial K_i}\right) \tag{14b}$$

The solution to eq 14 is

$$\begin{bmatrix} \Delta K_1 \\ \Delta K_2 \\ \cdots \\ \Delta K_n \end{bmatrix} = \begin{bmatrix} fK_1 \cdot fK_1 & fK_2 \cdot fK_1 & \cdots & fK_n \cdot fK_1 \\ fK_1 \cdot fK_2 & fK_2 \cdot fK_2 & \cdots & fK_n \cdot fK_2 \\ \cdots & \cdots & & \cdots \\ fK_1 \cdot fK_n & fK_2 \cdot fK_n & \cdots & fK_n \cdot fK_n \end{bmatrix}^{-1} \begin{bmatrix} D \cdot fK_1 \\ D \cdot fK_2 \\ \cdots \\ D \cdot fK_n \end{bmatrix} \tag{15}$$

The matrix requiring inversion in eq 15 is generally an ill-conditioned matrix. It is approximately equal to the principal minor of order $n + 1$ of the infinite Hilbert matrix.[19] This means that extreme care is required to obtain a reasonably accurate inversion. It has been found that attempting to invert the matrix of eq 15 by a standard Gauss–Jordan reduction is thoroughly unsatisfactory for this work. The inversion was, therefore, performed by what amounts to the long hand procedure of computing all n^2 elements directly. The resultant procedure was tested against the familiar Hilbert matrix. For the principle minor of degree eight, the minimax norm[20] of the product of the inverse and the starting matrix differed from unity by less than 10^{-11}.

Functional Constraints

One of the most powerful features of the analytical approach outlined above is the facility with which one may constrain the coefficients to a particular range or alter the basic functional form to provide a more sharply defined description of the particular system being studied. As a direct illustration, one of the very common problems encountered in the previous analyses of heterogeneous systems has been the occurrence of apparent negative concentrations.[21] This eventuality may be eliminated by redefining all of the $c_{i,\mathrm{ref}}$ coefficients as the square of a secondary parameter and evaluating this secondary term. The net result of this approach is to allow $c_{i,\mathrm{ref}}$ to range freely from 0 up through any positive value, but not to become negative. An identical constraint is also employed for the values of M_i where negative values are likewise physically unreal. The only complication involved in imposing such constraints is that care must be taken in developing the $(\partial f_r/\partial K_i)$ parameters in terms of the constrained coefficients.

This also allows the flexibility to distinguish between the case of concentration increments and of absolute concentrations. In those instances where the actual concentrations are known, experiments employing absorption optics or interference optics using either the method of Yphantis[1] or of Teller, *et al.*,[7] one need only set δ equal to 0 so that y_r equals c_r. For the major applications of the normal interference system, however, it is highly advantageous to be able to include the base increment as a determinable unknown.

It is possible by a minor modification in the development of eq 13 to treat directly the phenomenon of restricted self-association. Since M_i, in this instance, is equal to iM_1, the general case of self-association can be produced by this simple constraint. In addition, if any particular aggregates are known or assumed to be missing, it is possible to define the basic function so that the terms corresponding to these degrees of association will be omitted. The case of association may now be represented as

$$c_r = y_r - \delta = \sum_i c_{i,\mathrm{m}} e^{iAM_1(1-\bar{v}\rho)\Delta r^2} \tag{13a}$$

in which i takes only the allowed values.

In this same manner, it is a simple matter to modify eq 7 to accept the case of mixed association, *i.e.*

$$nA + mB \rightleftharpoons A_nB_m$$

Such systems have been considered at some length.[10] They can be regarded as a special case of three independent sedimenting species. Components A and B are unrelated and must be treated as simply polydispersity. Component A_nB_m is evaluated either as a simple third component or by constraining its molecular weight to be the sum of nA and mB. By means of such a constraint the analysis for all three species would be far more reliable than if they were to be regarded as three unique components. In fact, it is likely that two products, as well as two reactants, could be resolved.

The same analytical routine, again with only minor alterations, permits analyses based on the observation of concentration gradients. By differentiation of eq 7 with respect to r one obtains

$$\left(\frac{dc}{dr}\right)_r = \left(\frac{dy}{dr}\right)_r = \sum_i 2AM_i(1 - \bar{v}\rho)rc_{i,\mathrm{m}} e^{AM_i(1-\bar{v}\rho)\Delta r^2} \tag{16}$$

Equation 16 permits one to use the gradients, observed from either the schlieren system or the photoelectric scanner, in the place of y_r. Equation 16 does, in fact, have one less unknown in that there is no concern regarding an increment in concentration. In addition, the function has a higher sensitivity to the molecular weight. This should permit facility in determining M in spite of the higher uncertainty in the experimental measurement.

In addition to the foregoing detailed modifications it is evident that minor modifications of the basic function can account for several additional possibilities. If one assumes, for example, that the density is, to a first approximation, a linear function of the concentration of low molecular weight cosolutes, one may approximate the density by an expression of the form

$$\rho_r = \rho_{\mathrm{m}}(1 + kr^2) \tag{17}$$

in which ρ_{m} is the density at the meniscus, ρ_r is the

(19) A. Ralston, "A First Course in Numerical Analysis," McGraw-Hill, New York, N. Y., 1965, pp 232–233.

(20) The minimax norm, $\| \times \|_\infty$, is defined as the vector x_μ of the matrix for which $\| \times \|_\infty = \max|x_\mu|$ $(\mu = 1, 2, \cdots n)$.

(21) W. P. Reinhardt and P. G. Squire, *Biochim. Biophys. Acta*, **94**, 566 (1965).

density at r, and k is a first-order approximation constant. If the cosolute is well characterized, it may be possible to estimate k. Alternatively, k might stand as an additional unknown in the distribution equation. In any case one can accommodate a density variation in the analytical form by a simple modification of the function. It is likewise possible to incorporate the necessary functional rearrangements to include factors such as preferential interactions[9] or pressure effects[22] as well. Further work is under way regarding the problem of variable preferential interactions in associating systems.

Results and Discussion

General Systems. The fundamental analytical procedure has been developed without preliminary assumptions regarding the composition or nature of the system under investigation. The imposition of an upper limit on the number of components is placed on the system at the time of analysis along with the preliminary guesses for the concentrations at the meniscus and the molecular weights. For the systems discussed herein, concentration increments were employed in all cases since this clearly provides the more stringent test of the method. If absolute concentrations are available, it is obvious that one fewer coefficient need be determined and the fitting process will, accordingly, be more reliable.

The "data" analyzed in this paper were all computer generated so as to provide test systems with known answers. In order to simulate a real system it was necessary to select a series of "experimental" parameters. The choice of conditions was premised on the established performance capabilities of the interference optical system present in the Spinco analytical ultracentrifuge. A normal random number generator capable of producing $2^{35} - 1$ successive numbers without repetition was employed to simulate a random error of ± 0.03 fringe unit in each determination of y_i. After randomization, the generating alogrithm recorded on magnetic tape the simulated set of concentrations and radial loci. This isolated set was then used as the data for analysis.

While the current tendency has been to use short solution columns, approximately 3 mm or less, for sedimentation equilibrium, it was premised that a 6-mm column could be used. This would, admittedly, imply requiring a fourfold increase in the time to reach equilibrium. It also provided the added range of information to permit the more complex analyses. Systems with two sedimenting species were the limit for a 3-mm column. All values used for analysis were restricted to the region of the cell with concentrations above 0.03 fringe and with gradients of 250 fringes/cm or lower. Through this interval 150 concentrations were determined at equally spaced increments in r. While this may seem a large number, it must be noted that only one determination was made at each radial locus. It was observed empirically that with a fixed number of determinations, *i.e.*, 150, this approach allowed a better final analysis than the observation of concentration at 75 loci with each locus determined twice or for 50 loci determined three times. Therefore, the acquisition of 150 points is not out of line with reading patterns reported elsewhere.[9] Finally, for convenience in the generation of the simulated data the partial specific volumes were uniformally set at 0.75 cc/g and the density at 1 g/cc.

TABLE I
ANALYSIS OF HOMOGENEOUS SYSTEMS

c_0	Rpm	$c_{meniscus}$ Input	$c_{meniscus}$ Output	Molecular weight[a] $\times 10^{-5}$ Output
40.0	4,000	26.7	26.3	1.011
	6,000	15.1	15.1	1.000
	8,000	6.13	6.01	1.018
	10,000	1.67	1.65	1.003
4.0	4,000	2.67	2.56	1.030
	6,000	1.51	1.50	1.005
	8,000	6.13×10^{-1}	5.87×10^{-1}	1.025
	10,000	1.67×10^{-1}	1.63×10^{-1}	1.005
	12,000	2.96×10^{-2}	2.73×10^{-2}	1.018
0.4	4,000	2.67×10^{-1}	1.80×10^{-1}	1.290
	6,000	1.51×10^{-1}	1.38×10^{-1}	1.045
	8,000	6.13×10^{-2}	5.95×10^{-2}	1.010
	10,000	1.67×10^{-2}	1.47×10^{-2}	1.048
	12,000	2.96×10^{-3}	2.35×10^{-3}	1.076
	14,000	3.41×10^{-4}	2.39×10^{-3}	1.066

[a] The input value is 100,000.

The process of testing the analytical procedure was divided into distinct stages; homogeneous systems, two-component nonassociating systems, three-component nonassociating systems, monomer–dimer systems, dimer–trimer–tetramer systems, tetramer–pentamer–heptamer–octamer systems and monomer–dimer–tetramer–octamer systems. Excellent results were found for all but the tetramer–pentamer–heptamer–octamer system. For simplicity the general behavior of the results will also be broken into the distinct sections outlined above.

The results obtained for the homogeneous systems were good. Without introducing any assumptions regarding the coefficients to be determined, except that both the concentration and molecular weights were positive, values for the molecular weight were correct to within 1% for an initial concentration (c_0) of 4 fringes (f). The performance with c_0 values of 40, 4, and 0.4 f are illustrated in Table I with regard to both the return of molecular weights and meniscus concentrations. The ability of the analytical procedure to test the number of components was also demonstrated by the inability to force fit the data from a homogeneous system to a two-component model. While this analysis is impressively reliable, it is not very exciting as it was possible to treat the homogeneous system reasonably well before.

The results from the analysis of two component systems are exciting. The best results were obtained when the ratio of M_1 to M_2 was between 1:2 and 1:3.[23] Three systems were analyzed with M_1 equal to 100,000 and M_2 equal to 275,000. The first assumed equal

(22) G. Kegeles, S. Kaplan, and L. Rhodes, *Ann. N. Y. Acad. Sci.*, **164**, 183 (1969).

(23) Scholte[8] reports a similar observation in the analysis of molecular weight distributions.

TABLE II
TWO-COMPONENT NONASSOCIATING SYSTEMS

c_0			$c_{meniscus}$				Molecular weights[a] × 10^{-5}	
			1		2		Output	
1	2	Rpm	Input	Output	Input	Output		
1.0	3.0	8,000	1.53×10^{-1}	1.68×10^{-1}	5.93×10^{-3}	6.22×10^{-3}	0.948	2.740
		9,000	8.42×10^{-2}	8.68×10^{-2}	8.18×10^{-4}	8.53×10^{-4}	0.984	2.750
		10,000	4.17×10^{-2}	4.13×10^{-2}	8.47×10^{-5}	8.67×10^{-5}	1.000	2.748
		11,000	1.85×10^{-2}	1.84×10^{-2}	6.62×10^{-6}	7.13×10^{-6}	1.000	2.765
		12,000	7.39×10^{-3}	7.52×10^{-3}	3.92×10^{-7}	4.35×10^{-7}	0.995	2.764
		13,000	2.65×10^{-3}	2.77×10^{-3}	1.77×10^{-8}	1.98×10^{-8}	0.991	2.741
2.0	2.0	8,000	3.06×10^{-1}	3.15×10^{-1}	3.96×10^{-3}	4.13×10^{-3}	0.984	2.786
		9,000	1.69×10^{-1}	1.70×10^{-1}	5.45×10^{-4}	5.78×10^{-4}	0.995	2.757
		10,000	8.34×10^{-2}	8.28×10^{-2}	5.64×10^{-5}	5.81×10^{-5}	1.001	2.752
		11,000	3.70×10^{-2}	3.64×10^{-2}	4.41×10^{-6}	4.98×10^{-6}	1.003	2.775
		12,000	1.48×10^{-2}	1.45×10^{-2}	2.61×10^{-7}	3.00×10^{-7}	1.003	2.775
		13,000	5.29×10^{-3}	5.27×10^{-3}	1.18×10^{-8}	1.25×10^{-8}	1.001	2.754
		14,000	1.71×10^{-3}	1.71×10^{-3}	4.03×10^{-10}	4.61×10^{-10}	0.999	2.762
3.0	1.0	9,000	2.53×10^{-1}	2.53×10^{-1}	2.73×10^{-4}	2.82×10^{-4}	0.999	2.747
		10,000	1.25×10^{-1}	1.24×10^{-1}	2.82×10^{-5}	3.00×10^{-5}	1.001	2.764
		11,000	5.56×10^{-2}	5.50×10^{-2}	2.21×10^{-6}	2.26×10^{-6}	1.002	2.763
		12,000	2.22×10^{-2}	2.19×10^{-2}	1.31×10^{-7}	1.33×10^{-7}	1.003	2.766
		13,000	7.94×10^{-3}	7.86×10^{-3}	5.89×10^{-9}	6.33×10^{-9}	1.002	2.770
		14,000	2.56×10^{-3}	2.54×10^{-3}	2.02×10^{-10}	2.61×10^{-10}	1.001	2.755
		15,000	7.42×10^{-4}	7.43×10^{-4}	5.27×10^{-11}	1.06×10^{-10}	1.000	2.784

[a] The input values for the molecular weights are 100,000 and 275,000.

TABLE III
EFFECT OF MOLECULAR WEIGHT RATIOS IN TWO-COMPONENT SYSTEMS

Second component	Component	c_0	$c_{meniscus}$		Molecular weights[a] × 10^{-5}
			Input	Output	Output
175,000	1	1	8.42×10^{-2}	8.80×10^{-2}	1.020
	2	3	2.43×10^{-2}	2.25×10^{-2}	1.830
	1	2	3.70×10^{-2}	3.72×10^{-2}	0.999
	2	2	8.73×10^{-3}	7.93×10^{-3}	1.800
	1	3	2.21×10^{-2}	2.21×10^{-2}	1.000
	2	1	7.70×10^{-5}	7.22×10^{-5}	1.780
150,000	1	2	3.70×10^{-2}	3.81×10^{-2}	0.990
	2	2	3.14×10^{-3}	2.97×10^{-3}	1.520
130,000	1	2	3.70×10^{-2}	3.61×10^{-2}	0.972
	2	2	8.59×10^{-3}	1.05×10^{-2}	1.290
120,000	1	2	3.70×10^{-2}	3.52×10^{-2}	0.961
	2	2	1.41×10^{-2}	1.72×10^{-2}	1.212

[a] The input value for the molecular weight of component 1 is 100,000.

values of c_0 (2.0 f), the second set c_{01} at 1.0 f and c_{02} at 3.0 f, and the third put c_{01} at 3.0 f and c_{02} at 1.0 f. The results for these systems are presented in Table II. It is obvious from the values in Table II that the quality of the analysis varies with the speed or more meaningfully with the nature of the distribution. The observed values for the molecular weights go through a maximum as the angular velocity is increased and this maximum clearly corresponds to the best value for the molecular weight. Although the values determined on the low side vary by several per cent, the overestimates are within 0.3% and indicate the appropriate set of coefficients quite reliably.

While a one-component system will work for any molecular weight, it is important to test the resolving power of the system for more complex combinations. In Table III we see the results obtainable for a series of ratios of M_2 to M_1. It is clear that the greater this ratio the better the analysis. However, even at a ratio of 1.2:1.0 the results are still significant.

The results from three-component systems were substantially more complex. Samples were observed as a function of concentration ratio as well as molecular weight variations. When the molecular weights were held reasonably close, 100,000:180,000:330,000, results were obtained which were useful, but subject to an uncertainty as much as 10%. There is, again, a substantial dependence on velocity. The values approach the correct range monotonically. In most instances, however, the correct values are marked by the fact that attempts to obtain still greater redistributions of the macromolecules produce a divergence in the analysis. The concentration coefficients in these analyses are of limited significance. In Table IV we see the results obtained at the speed of choice for each of several c_0 combinations. The results indicated are, in every case,

TABLE IV
ANALYSIS OF THREE-COMPONENT NONASSOCIATING SYSTEMS

Component[a]	c_0	$c_{meniscus}$ Input	$c_{meniscus}$ Output	Molecular weight × 10^{-5} Output
1	0.5	3.69×10^{-3}	5.01×10^{-3}	0.950
2	1.0	5.63×10^{-5}	2.80×10^{-5}	1.94
3	1.5	5.50×10^{-9}	1.95×10^{-9}	3.46
1	0.7	5.97×10^{-4}	6.37×10^{-4}	1.00
2	1.0	9.04×10^{-7}	1.32×10^{-7}	1.98
3	1.3	1.89×10^{-12}	2.64×10^{-13}	3.54
1	1.0	2.47×10^{-4}	2.48×10^{-4}	1.007
2	1.0	8.73×10^{-8}	2.74×10^{-9}	2.10
3	1.0	1.79×10^{-14}	2.08×10^{-15}	3.54
1	1.3	8.41×10^{-5}	9.54×10^{-5}	0.988
2	1.0	7.04×10^{-9}	2.46×10^{-9}	1.91
3	0.7	1.11×10^{-16}	1.32×10^{-18}	3.59
1	1.5	3.71×10^{-4}	5.33×10^{-4}	0.942
2	1.0	8.73×10^{-8}	5.55×10^{-7}	1.64
3	0.5	8.93×10^{-15}	3.73×10^{-15}	3.46
1	1.0	2.65×10^{-3}	2.87×10^{-3}	1.013
2	4.0	3.13×10^{-5}	9.66×10^{-6}	1.99
3	7.0	6.02×10^{-10}	7.76×10^{-11}	3.56
1	2.0	1.29×10^{-4}	1.19×10^{-4}	1.02
2	4.0	2.82×10^{-8}	1.19×10^{-9}	2.12
3	6.0	9.52×10^{-16}	5.27×10^{-17}	3.68
1	4.0	6.11×10^{-5}	7.17×10^{-5}	0.983
2	4.0	1.90×10^{-9}	1.67×10^{-9}	1.83
3	4.0	4.09×10^{-18}	1.63×10^{-18}	3.54
1	6.0	1.48×10^{-3}	2.26×10^{-3}	0.888
2	4.0	3.49×10^{-7}	5.85×10^{-5}	1.32
3	2.0	3.57×10^{-14}	2.02×10^{-13}	3.45
1	7.0	1.73×10^{-3}	2.85×10^{-3}	0.95
2	4.0	3.49×10^{-7}	2.99×10^{-6}	1.41
3	1.0	1.78×10^{-14}	1.52×10^{-9}	3.50

[a] Component 1, 100,000; component 2, 180,000; component 3, 330,000.

the maximum molecular weight values reached with increasing speed. In some cases these values were followed by lower results as the speed increased while in others the analysis either diverged or converged on two components instead of three. Although neither the concentrations nor the molecular weight values are comparable to what one expects for a pure sample, they are unquestionably reasonably good values.

Table V illustrates the results obtained with some modifications of conditions for the three component system. It is demonstrated that a broad range of molecular weight ratios can be dealt with while maintaining sufficient accuracy to be useful. The use of the higher minimum reading was explored as a possible factor which obviously can be significant. Where the concentrations at the meniscus are so low that the flatness of the fringe pattern will clearly imply a near to zero value, there is no advantage to obtaining readings at very slight displacements.

Associating Systems. The two major features of the algorithm employed for the analysis of associating systems are the ability to select specific stages of association to be allowed while excluding all others and the requirement that all of the molecular weights employed are integral multiples of a common monomer.

TABLE V
ANALYSIS OF THREE-COMPONENT NONASSOCIATING SYSTEMS

Component	c_0	$c_{meniscus}$ Input	$c_{meniscus}$ Output	Molecular weight × 10^{-5} Output
1[a]	1.0	2.47×10^{-4}	2.57×10^{-4}	1.00
2[b]	4.0	2.76×10^{-8}	1.17×10^{-8}	2.15
3[c]	7.0	6.36×10^{-18}	4.30×10^{-19}	4.56
1[a]	2.0	4.95×10^{-4}	7.65×10^{-4}	0.93
2[b]	4.0	2.76×10^{-8}	2.19×10^{-7}	1.84
3[c]	6.0	5.45×10^{-18}	6.03×10^{-18}	4.42
1[a]	4.0	9.90×10^{-4}	1.49×10^{-3}	0.923
2[b]	4.0	2.76×10^{-8}	2.81×10^{-6}	1.59
3[c]	4.0	3.64×10^{-18}	2.64×10^{-17}	4.29
1[a]	4.0	3.41×10^{-3}	3.76×10^{-3}	0.983[f]
2[b]	4.0	4.03×10^{-7}	5.87×10^{-7}	2.02
3[c]	4.0	1.10×10^{-15}	1.27×10^{-15}	4.51
1[a]	6.0	5.12×10^{-3}	5.86×10^{-3}	0.972
2[b]	4.0	4.03×10^{-7}	2.86×10^{-6}	1.83
3[c]	2.0	5.48×10^{-16}	2.17×10^{-15}	4.48
1[a]	1.0	2.47×10^{-4}	2.58×10^{-4}	1.003
2[d]	1.0	8.73×10^{-8}	1.96×10^{-9}	2.171
3[e]	1.0	5.27×10^{-12}	2.32×10^{-13}	3.086

[a] 100,000. [b] 205,000. [c] 425,000. [d] 180,000. [e] 275,000. [f] "Data" were accepted only for concentrations above 0.1 fringe.

The four basic systems considered are represented in Table VI.

The first system is a monomer–dimer system. The molecular weight returned for the monomer is well within reasonable limits. The equilibrium constant is quite good being within 7% of the input value. The second is also straightforward, a dimer–trimer–tetramer system. The molecular weight values are again quite reliable. This group illustrates the power of being able to stipulate the composition since the monomer molecular weight was computed even though it was not allowed to exist in the system. Not only was the molecular weight returned to within 2% but the respective concentrations were adequate to obtain noteworthy approximations for the equilibrium constants. For convenience, these constants are computed in terms of a stoichiometric reaction converting the lowest molecular weight species present to the aggregate of interest.

The other three systems in Table VI are substantially more complex. Each of these utilizes the facility for selecting the stages of aggregation to be computed. However, one additional species is assumed to be present to test the potential of the program to operate when a complete definition of the components is not available. The first set is by far the most intractable in the sensitivity and resolution required for analysis. It represents data from a tetramer–pentamer–heptamer–octamer system, but does not exclude the hexamer from the analysis. In spite of this the monomer molecular weight is returned with only a 12% error. The concentrations are, however, not exactly spectacular.

The final two systems of Table VI illustrate very powerfully the versatility of this procedure. The data for both of these sets have been generated for a monomer–dimer–tetramer–octamer system. In the analysis

TABLE VI
ANALYSIS OF ASSOCIATING SYSTEMS

Components considered	Component no.	$c_{meniscus}$ Input	$c_{meniscus}$ Output	Monomer molecular weight × 10^{-5} Output	K_a[a] Input	K_a[a] Output
100,000	1	4.17×10^{-2}	4.03×10^{-2}	1.015		
200,000	2	7.20×10^{-4}	6.13×10^{-4}		0.41	0.38
100,000	2	1.85×10^{-2}	1.82×10^{-2}	0.508		
150,000	3	1.57×10^{-3}	1.38×10^{-3}		0.39	0.32
200,000	4	1.19×10^{-4}	9.78×10^{-5}		0.35	0.30
100,000	2	1.53×10^{-1}	1.57×10^{-1}	0.509		
150,000	3	4.85×10^{-2}	5.20×10^{-2}		0.66	0.64
200,000	4	1.40×10^{-2}	8.76×10^{-3}		0.60	0.36
100,000	4	4.17×10^{-2}	9.8×10^{-2}	0.280		
125,000	5	1.58×10^{-2}	3.0×10^{-1}			
150,000	6	0.0	5.0×10^{-2}			
175,000	7	2.02×10^{-3}	1.2×10^{-3}			
200,000	8	7.20×10^{-4}	7.5×10^{-3}			
25,000	1	4.49×10^{-1}	1.5	0.26		
50,000	2	1.73×10^{-1}	3.1×10^{-3}			
100,000	4	1.85×10^{-2}	2.5×10^{-1}			
150,000	6	0.0	4.9×10^{-4}			
200,000	8	1.19×10^{-4}	4.5×10^{-4}			
25,000	1	3.12×10^{-1}	3.25×10^{-1}	0.256		
50,000	2	7.42×10^{-2}	6.39×10^{-2}		0.76	0.60
100,000	4	2.65×10^{-3}	2.36×10^{-3}		0.28	0.21
150,000	6	0.0	2.57×10^{-8}		0.0	2.2×10^{-5}
200,000	8	1.75×10^{-6}	1.34×10^{-6}		1.98×10^{-2}	1.08×10^{-2}

[a] The values of K_a are all computed in terms of the separate stoichiometric conversion of the lowest molecular weight species present to the aggregate of interest.

it has been assumed that the absence of trimer, pentamer, and heptamer were known but that the hexamer was uncertain. By means of the general solution it was possible to obtain a molecular weight high by only 4%; however, the concentration values at the meniscus were of very limited quality. A minor alteration in the "data acquisition" produced a dramatic change. Data were considered only for concentrations above 0.1 f instead of 0.03 f as had been done previously. The molecular weight value came to within 2.5% of 25,000; the absence of a hexamer became clear. Even more the association constants were now usably close to the proper values.

Summary

It is evident that the analytical procedure developed is capable of treating comparatively complex mixtures at sedimentation equilibrium. Not only are reliable values for the molecular weights obtained, but the concentrations of the various components are returned with sufficient accuracy to permit the determination of equilibrium constants for associating systems. Fortran V listings (Univac 1108 Fortran code) are available from the author. The analysis of associating systems in complex solvents where the phenomenon of preferential interaction occurs is currently being developed and will be reported in a subsequent paper.