Enumeration of Species in a Self-Associating Protein System by Diffusion Measurements in the Ultracentrifuge*

Wilhelm Godschalk

ABSTRACT: A method is proposed for the enumeration o macromolecular species in a self-associating system. The experimental technique consists of the formation of a boundary in an analytical ultracentrifuge cell, followed by rapid deceleration to a low rotor speed, where sedimentation is negligible. The boundary spreading is observed for several hours, and concentration profiles are recorded at fixed time intervals, by means of ultraviolet absorption optics. For a complete analysis of an interacting system, several experiments, at different bulk concentrations, are necessary. The concentration profiles are subjected to an integral transformation of the finite Hankel type. The resulting transforms are arranged in a matrix, the rows of which represent the values obtained at different time intervals, for one particular sample. A rank determination on this matrix yields the number of interacting species in the system. It is shown that the matrix is the product of two other matrices, one of which is an array of initial values of the transform, the other being composed entirely of pure time-dependent functions containing the diffusion coefficients for the system. A similar approach could be used for conventional diffusion measurements in a longer column. Although this would give rise to data of higher accuracy, the equipment would be quite specialized and far less generally available than the analytical ultracentrifuge with uv scanning.

As a result of the growing recognition of the importance of cooperative phenomena in biochemistry, the study of multisubunit structures has been a subject of considerable interest for several years. One of the most widely used techniques for studying subunit aggregation is equilibrium sedimentation. Although significant progress has been made in recent years, evaluation of molecular weight distributions is still a difficult problem, especially when higher polymers are present. The usual curve-fitting approach can be most effective for disproving a proposed model; however, consistence of a model with the available experimental data still leaves the question of whether this model is unique.

An independent determination of the number of species, participating in the association equilibria, would provide very valuable additional information. An elegant method has been developed by Ackers (1968), using molecular sieve chromatography. The underlying principle is, that the spectrum of weight-average partition coefficients, each pertaining to an equilibrium state of the reacting system and a different gel porosity, defines a linear vector space. The dimension of the space obtained by matrix rank analysis yields the number of interacting components. This argument leads to the important conclusion that it must be possible and advantageous to apply the methods of spectral analysis to transport phenomena. As Ackers points out, such an attempt can only be successful if a parameter can be found which varies in a linearly independent fashion with molecular size.

Spectral analysis itself can be considered a classical concept, and presents one of the more beautiful generalizations of mathematics (*cf.* Kharkevich, 1960).

The minimum number of components in a system can be found by sampling a spectral quantity over a range of concentrations, arranging the data in a matrix, and performing a matrix rank analysis (Wallace, 1960; Weber, 1961; Ainsworth, 1961; Wallace and Katz, 1964; McMullen *et al.*, 1967; Ainsworth and Bingham, 1968).

Ackers' molecular sieve method is most powerful with the newly developed ultraviolet scanning system (Brumbaugh and Ackers, 1968); the equipment, however, is not yet commonly available. The method described in this paper is a novel approach to a known centrifugal technique, using commercial equipment available in most laboratories. The experimental part consists of the formation of a boundary in an analytical ultracentrifuge cell, followed by rapid deceleration to the lowest speed which the instrument will allow without rotor precession. The name "field relaxation" for this procedure has been introduced by Kegeles and Sia (1963). After deceleration, the boundary spreading is observed for several hours, and concentration profiles are recorded at fixed time intervals, by means of ultraviolet absorption optics. A similar approach has been used previously for the determination of diffusion coefficients (Markham, 1962, 1967). Indeed, the diffusion process governs the flow of solute during the experiment; the sedimentation force is negligible at the low angular speeds used. The main feature of the proposed procedure is the data analysis. A new characteristic parameter, which is a nonlinear function of the diffusion coefficient, constitutes a typical element of the matrix to be subjected to rank determination. This parameter is based upon an integral transformation, and will be defined in the next section. In this and future papers the proposed method will be referred to as the *diffusion boundary transformation method.*

The experimental feasibility of the procedure has been tested on α-chymotrypsin, which is known to be a monomer–dimer system at low pH (Winzor and Scheraga, 1964). However, the experimental results will be reported in the next paper of this series.

Theory

Single Solute Species. The theory of diffusion-controlled flow of a single macromolecular species in a sector-shaped

* From the Department of Biochemistry, University of Virginia School of Medicine, Charlottesville, Virginia 22901. *Received February 1, 1971.* This work was supported by U. S. Public Health Service Grant GM-12569. This paper is the first in a series on measurements of the molecular parameters of interacting protein systems in the ultracentrifuge.

centrifuge cell, after rapid deceleration, has been presented in a previous paper (Godschalk, 1968). Only the results pertinent to the present discussion will be recalled.

For a single, ideal component, an integral transform of the concentration distribution is defined as

$$\tilde{C}_j \equiv \int_a^b C_j B_0(pr) r dr \tag{1}$$

a and b are the radii of the meniscus and bottom of the liquid column, respectively. The function $B_0(pr)$ is defined as

$$B_0(pr) \equiv J_0(pr) Y_0'(pa) - Y_0(pr) J_0'(pa) \tag{2}$$

in which J_0 and Y_0 are Bessel functions of zero order and of the first and second kind, respectively. Furthermore

$$J_0'(pa) \equiv \left[\frac{\partial J_0(pr)}{\partial(pr)}\right]_{r=a} \qquad Y_0'(pa) \equiv \left[\frac{\partial Y_0(pr)}{\partial(pr)}\right]_{r=a} \tag{3}$$

The values of the parameter, p, are given by the positive roots of the equation

$$J_0'(pb) Y_0'(pa) - Y_0'(pb) J_0'(pa) = 0 \tag{4}$$

with $J_0'(pb)$ and $Y_0'(pb)$ defined as in eq 3, but for $r = b$.

Under conditions of negligible sedimentation, the flow in the cell can be described by Fick's second law for cylindrical diffusion, and the boundary conditions consist of the requirement that the concentration gradient be zero at both $r = a$ and $r = b$. Application of this cylindrical transformation to the diffusion problem leads to the simple expression (Godschalk, 1968)

$$\tilde{C}_j = \overline{C_j^0} e^{-D_j p^2 t} \tag{5}$$

in which $\overline{C_j^0}$ represents the value of $\tilde{C}_j$ at time $t = 0$, and D_j is the diffusion coefficient of species j.

The quantities $\tilde{C}_j$ and $\overline{C_j^0}$ can only be determined experimentally in the case of a single, ideally behaving macromolecular species in solution. For such a system, a logarithmic plot of $\tilde{C}_j$ *vs.* time will have a slope of $-D_j p^2$. As p is known from the dimensions of the cell, and may be any positive root of eq 4, one centrifuge run is sufficient for the determination of the diffusion coefficient of a single macromolecular component.

Noninteracting Multicomponent System. For a mixture of macromolecular species, direct determination of individual diffusion coefficients is, of course, impossible. In this case the experiment yields average quantities, and a number of runs must be performed for a complete analysis.

Flow equations for multicomponent systems can be found in many forms, throughout the literature (*e.g.*, Fujita, 1962). For our purpose, it is advantageous to start from the basic premises of nonequilibrium thermodynamics (*cf.* Katchalsky and Curran, 1965; De Groot and Mazur, 1962), so that the theoretical treatment remains uniform for all present and future cases to be discussed.

For a system containing n solute species in a common solvent, the flows of the individual solute species are given in terms of the phenomenological equations

$$J_k = -\sum_{j=1}^{n} L_{kj} \operatorname{grad} \mu_j \qquad (k = 1, 2, \ldots n) \tag{6}$$

In this notation, J_k is the flow of species k, μ_j is the chemical potential of species j, and L_{kj} is a phenomenological coefficient. The frame of reference is the *volume-fixed frame.* This frame, however, coincides with the cell-fixed frame, provided the partial specific volume of each component is constant, which we will assume in the following analysis.

A general expression for the gradients of the chemical potentials can be written as

$$\operatorname{grad} \mu_j = \sum_{i=1}^{n} \frac{\partial \mu_j}{\partial C_i} \operatorname{grad} C_i \qquad (j = 1, 2, \ldots n) \tag{7}$$

So that

$$J_k = -\sum_{i=1}^{n} \sum_{j=1}^{n} L_{kj} \frac{\partial \mu_j}{\partial C_i} \operatorname{grad} C_i \qquad (k = 1, 2, \ldots n) \tag{8}$$

If we compare this expression with the customary diffusional flow equations in terms of the set of diffusion coefficients, D_{ki}, we can formally equate these coefficients with the parameters of eq 8

$$D_{ki} = \sum_{j=1}^{n} L_{kj} \frac{\partial \mu_j}{\partial C_i} \qquad \begin{matrix}(k = 1, 2, \ldots n) \\ (i = 1, 2, \ldots n)\end{matrix} \tag{9}$$

These equations will serve as background information on the nature of the coefficients D_{ki}. Although the general form will be used in the treatment of the nonideal case, to be reported in a future communication, we restrict the present discussion to the behavior of ideal solutions. As the chemical potential of a species j in a thermodynamically ideal system can be written as

$$\mu_j = \mu_j^0 + RT \ln C_j \tag{10}$$

all derivatives $\partial \mu_j / \partial C_i$ with $j \neq i$ are zero. If the system is also hydrodynamically ideal, *i.e.*, no coupling exists between the flows, then the cross coefficients, $L_{kj} (k \neq j)$ can be ignored. The combination of thermodynamic and hydrodynamic ideality causes the cross-diffusion coefficients to vanish, provided scalar flows (chemical reactions) are absent. All these restrictions simplify the flow equations considerably; we simply have now

$$J_k = -D_{kk} \operatorname{grad} C_k \qquad (k = 1, 2, \ldots n) \tag{11}$$

For the sector-shaped ultracentrifuge cell, this leads to the continuity equation

$$\frac{\partial C_k}{\partial t} = D_{kk} \frac{\partial^2 C_k}{\partial r^2} + \frac{D_{kk}}{r} \frac{\partial C_k}{\partial r} \qquad (k = 1, 2, \ldots n) \tag{12}$$

It has been assumed that the coefficients D_{kk} are independent of r. For dilute solutions, this is not an unreasonable assumption. Moreover, it has been shown experimentally that the "straight" phenomenological coefficients, L_{kk}, are, to a good approximation, linearly dependent on the concentration of the components to which they correspond (*cf.* Fujita and Gosting, 1960). As $D_{kk} = L_{kk} \partial \mu_k / \partial C_k$, and for an ideal system $\partial \mu_k / \partial C_k = RT/C_k$, it follows that D_{kk} is approximately constant.

Not surprisingly, eq 12 has precisely the form of Fick's second law for a single solute component. Thus, eq 12 ex-

presses a conclusion which could have been drawn intuitively. If no interactions between the flows of the different species occur, then each species diffuses as if it were present alone in solution.

Equation 12 can now be transformed in the same fashion as the expression for a single species, by multiplication with $B_0(pr)r$, and integration over the range a to b. The result is completely analogous to eq 5

$$\bar{C}_k = \overline{C_k^0} e^{-D_{kk}p^2t} \qquad (k = 1, 2, \ldots n) \tag{13}$$

If we define the transform of the total concentration distribution as

$$\bar{C} \equiv \int_a^b C B_0(pr) r \mathrm{d}r \tag{14}$$

in which C is the total concentration at each point, then we can write

$$\bar{C} = \sum_{k=1}^{n} \overline{C_k^0} e^{-D_{kk}p^2t} \tag{15}$$

This is a consequence of the linearity of the transformation; the transform of the sum of the individual concentrations, C_k, equals the sum of the transforms $\bar{C}_k$.

Self-Associating System. The flow equations for interacting systems can be developed along the same lines as those of the previous section. Equations 6–9 are perfectly valid for this case. The important difference is that, for the interacting system, a chemical coupling exists between the concentration gradients; this leads to a coupling between the flows. Of the many possible systems with chemical interactions, we shall examine one which is of great interest in biochemistry; a self-associating macromolecule in solution. It will be assumed that the rate of the association reactions is sufficiently fast compared to the diffusion process, so that chemical equilibrium is attained at each point in the centrifuge cell.

Starting with eq 9, we can observe that, even if all activity coefficients are unity, so that eq 10 holds, and all coefficients $L_{kj}(k = j)$ are ignored, the cross-diffusion coefficients will not vanish.

To avoid the burdensome formalism of the summations, a concrete example will be given which can be readily expanded to the more general case. Let us consider a solution containing monomer, dimer, and trimer in chemical equilibrium. This system is characterized by nine diffusion coefficients; these are, however, not all independent. There are six relations between them, which reduces the number of independent coefficients to three. Using eq 9, and ignoring all cross-term phenomenological coefficients, we can tabulate the D's as follows

$$\begin{aligned}
D_{11} &= L_{11}\frac{\partial \mu_1}{\partial C_1} & D_{12} &= L_{11}\frac{\partial \mu_1}{\partial C_2} & D_{13} &= L_{11}\frac{\partial \mu_1}{\partial C_3} \\
D_{21} &= L_{22}\frac{\partial \mu_2}{\partial C_1} & D_{22} &= L_{22}\frac{\partial \mu_2}{\partial C_2} & D_{23} &= L_{22}\frac{\partial \mu_2}{\partial C_3} \\
D_{31} &= L_{33}\frac{\partial \mu_3}{\partial C_1} & D_{32} &= L_{33}\frac{\partial \mu_3}{\partial C_2} & D_{33} &= L_{33}\frac{\partial \mu_3}{\partial C_3}
\end{aligned} \tag{16}$$

From eq 10 it will be obvious that, in general

$$\frac{\partial \mu_j}{\partial C_k} = \frac{RT}{C_j}\frac{\partial C_j}{\partial C_k} \tag{17}$$

If the equilibrium constants for dimerization and trimerization are K_2 and K_3, respectively, it follows that

$$\begin{aligned}
D_{11} &= \frac{RTL_{11}}{C_1} & D_{12} &= \frac{RTL_{11}}{2K_2C_1^2} & D_{13} &= \frac{RTL_{11}}{3K_3C_1^3} \\
D_{21} &= \frac{2RTL_{22}}{C_1} & D_{22} &= \frac{RTL_{22}}{K_2C_1^2} & D_{23} &= \frac{2RTL_{22}}{3K_3C_1^3} \\
D_{31} &= \frac{3RTL_{33}}{C_1} & D_{32} &= \frac{3RTL_{33}}{2K_2C_1^2} & D_{33} &= \frac{RTL_{33}}{K_3C_1^3}
\end{aligned} \tag{18}$$

If we choose the straight coefficients, D_{11}, D_{22}, and D_{33} to characterize this system, all other coefficients are related to these three by the set of eq 18. Since, for this particular system, the gradients are related in the following way

$$\begin{aligned}
\text{grad } C_2 &= 2K_2C_1 \text{ grad } C_1 \\
\text{grad } C_3 &= 3K_3C_1^2 \text{ grad } C_1
\end{aligned} \tag{19}$$

we can write the flows as

$$\begin{aligned}
J_1 &= -D_{11} \text{ grad } C_1 - D_{12} \text{ grad } C_2 - D_{13} \text{ grad } C_3 = -3D_{11} \text{ grad } C_1 \\
J_2 &= -D_{21} \text{ grad } C_1 - D_{22} \text{ grad } C_2 - D_{23} \text{ grad } C_3 = -3D_{22} \text{ grad } C_2 \\
J_3 &= -D_{31} \text{ grad } C_1 - D_{32} \text{ grad } C_2 - D_{33} \text{ grad } C_3 = -3D_{33} \text{ grad } C_3
\end{aligned} \tag{20}$$

Treating D_{11}, D_{22}, and D_{33} as independent of r again, the continuity equations for the centrifuge become

$$\frac{\partial C_k}{\partial t} = 3D_{kk}\frac{\partial^2 C_k}{\partial r^2} + \frac{3D_{kk}}{r}\frac{\partial C_k}{\partial r} \qquad (k = 1, 2, \ldots n) \tag{21}$$

Each equation in this set can be transformed in the usual way, and the three individual transforms can be summed to give the transform of the total concentration

$$\bar{C} = \overline{C_1^0} e^{-3D_{11}p^2t} + \overline{C_2^0} e^{-3D_{22}p^2t} + \overline{C_3^0} e^{-3D_{33}p^2t} \tag{22}$$

If we compare this result with eq 15, for a noninteracting system, it is obvious that the principal difference is the factor 3 in the exponent. This makes the exponentials approach zero faster with time, than is the case with eq 15. This result is in agreement with the experimental fact that a boundary diffuses out faster if the various species interact chemically. The expression for the general case of a self-associating system has a factor n in the exponent, and the sum consists of n terms. For large values of t, the transform $\bar{C}$ vanishes. This is consistent with the fact that at infinite time the concentration becomes uniform, while the integral

$$\int_a^b B_0(pr) r \mathrm{d}r$$

equals zero (Godschalk, 1968).

It should also be noted that the initial values, $\overline{C_k^0}$ in eq 22 are interdependent, albeit not in a simple manner. The $\overline{C_k^0}$ of eq 15, on the other hand, are not related, except for the fact that they should add up to $\overline{C^0}$.

Thus, we can formally write the transform of the total concentration as the sum of the transforms of the individual species. The function $\overline{C_k}^0$ is dependent on the initial concentration distribution; the exponential, for which I propose the name *decay function*, is a function of time only, and is independent of the initial concentration, or the previous history of the experiment. The transform, $\bar{C}$, is thus composed of n products of a pure time function and a time-independent function of the initial distribution. It can be shown (W. Godschalk, unpublished data) that equations of this same general form result when nonideality is allowed; the only difference will be in the exponent of the decay function. In the most extreme case, the exponent will not be quite linear in t; even then, the general method set forth in the next sections of this paper, is still applicable.

Setting up the Matrix. Let us assume that we subject a total number of m samples, at different bulk concentrations, C_s^0, to field relaxation runs, and that concentration profiles are recorded at q time intervals, which must be equal for each run. If we write the symbol θ for the decay function, e^{-nDkkp^2t}, then a set of $m \times q$ equations of the following type can be set up

$$\sum_{k=1}^{n} \overline{C_{sk}}^0 \theta_{kt} = \bar{C}_{st} \qquad \begin{matrix}(s = 1, 2, \ldots . m)\\(t = 1, 2, \ldots . q)\end{matrix} \tag{23}$$

The summation is over all solute species present. In addition to these equations, there are m relations of the type

$$\sum_{k=1}^{n} \overline{C_{sk}}^0 = \overline{C_s}^0 \qquad (s = 1, 2, \ldots . m) \tag{24}$$

We can write the set of eq 23 in matrix form

$$\begin{bmatrix} \overline{C_{11}}^0 & \overline{C_{12}}^0 \ldots \overline{C_{1k}}^0 \ldots \overline{C_{1n}}^0 \\ \overline{C_{21}}^0 & \vdots \quad \vdots \\ \vdots & \vdots \quad \vdots \\ \overline{C_{s1}}^0 \ldots \ldots \overline{C_{sk}}^0 & \overline{C_{sn}}^0 \\ \vdots & \vdots \quad \vdots \\ \overline{C_{m1}}^0 \ldots \ldots \overline{C_{mk}}^0 \ldots \overline{C_{mn}}^0 \end{bmatrix} \begin{bmatrix} \theta_{11} & \theta_{12} \ldots \theta_{1t} \ldots \theta_{1q} \\ \theta_{21} & \vdots \quad \vdots \\ \vdots & \vdots \quad \vdots \\ \theta_{k1} \ldots \ldots \theta_{kt} \ldots \theta_{kq} \\ \vdots & \vdots \quad \vdots \\ \theta_{n1} \ldots \ldots \theta_{nt} \ldots \theta_{nq} \end{bmatrix} = \begin{bmatrix} \bar{C}_{11} & \bar{C}_{12} \ldots \bar{C}_{1t} \ldots \bar{C}_{1q} \\ \bar{C}_{21} & \vdots \quad \vdots \\ \vdots & \vdots \quad \vdots \\ \bar{C}_{s1} \ldots \ldots \bar{C}_{st} \ldots \bar{C}_{sq} \\ \vdots & \vdots \quad \vdots \\ \bar{C}_{m1} \ldots \ldots \bar{C}_{mt} \ldots \bar{C}_{mq} \end{bmatrix} \tag{25}$$

or, abbreviated

$$[\overline{C_{sk}}^0][\theta_{kt}] = [\bar{C}_{st}] \tag{26}$$

The matrix on the right in eq 25 contains the experimentally obtainable total transforms. This $m \times q$ matrix can be subjected to rank analysis; if m is sufficiently large (*i.e.*, enough runs are performed), the rank found will be equal to the number of molecular species in solution. Once the total number of species is known, it is, at least theoretically, possible to solve the whole system of equations for all $\overline{C_{sk}}^0 - \theta_{kt}$ pairs. In general, it will be advantageous to overdetermine the system, in order to reduce experimental errors. At this point, however, we are still concerned exclusively with the enumeration of species.

Matrix Rank Analysis

As was pointed out in the introduction, methods and applications of rank determination have been described previously by several authors. The methods are generally based on the theorem that no finite sequence of elementary transformations can alter the rank of a matrix. As the purpose of this paper is to present the complete procedure for the diffusion boundary transformation method, a brief description of the technique of rank analysis follows.

Consider the matrix

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \ldots . \\ a_{21} & a_{22} & a_{23} \ldots . \\ a_{31} & a_{32} & a_{33} \ldots . \\ \vdots & & \end{bmatrix}$$

In order to arrive at the desired result, a reduction process must be carried out, consisting of simple arithmetic operations, aimed at making as many elements as possible equal to zero. All rows are normalized with respect to the first row by the simple transformation

$$a_{ij}' = a_{ij} - \frac{a_{i1}}{a_{11}} a_j$$

i.e., a new second row is found by multiplication of the first row by $-a_{21}/a_{11}$ and adding the result to the original second row. Likewise, the old third row is added to the product of the first row and the factor $-a_{31}/a_{11}$, etc. The process is applied to all the remaining rows of the matrix. Now all elements, a_{i1}, of the first column, with the exception of a_{11}, will be zero. If all other elements, except those of the first row, which have remained unchanged, have been reduced to zero by this first reduction step, then the rank of the matrix is equal to one. If this condition is not met, a second reduction step is performed, on the submatrix obtained by omitting the first row and the first column of the original matrix. This procedure is repeated until at some stage all elements of the submatrix are zero. The rank will always be equal to the number of reduction steps necessary to reach this condition. Obviously, one should have a fairly accurate knowledge of the magnitude of the experimental errors associated with the elements of the original matrix; in order to arrive at the right conclusion, a zero test must be set up, based on these errors. Whenever a matrix element, after a reduction step, falls below the error value, it is considered zero. Propagation of errors through the reduction process can be a serious problem, unless sufficient care is exercised to minimize it. This aspect was given due consideration by Wallace and Katz (1964); these authors recommend interchanging rows and columns to the effect of making a_{11} the largest element in the matrix. A slightly different procedure for the minimization of error propagation has been developed in this laboratory, and will be reported elsewhere.

Detailed Procedure

(1) A boundary is formed at fairly high speeds (10,000 rpm or more), by layering 200 μl of solvent (buffer) on top of 400 μl of the protein solution, in a synthetic boundary cell. If the ultraviolet scanner of the Spinco Model E centrifuge is used, up to five samples, at different concentrations, can be used with the multiplexer. The commercially available double-sector–synthetic boundary centerpieces are perfectly suita-

ble for this purpose. (2) With the brake in the "rapid" position, the rotor speed is reduced to the lowest speed at which the rotor does not wobble. This speed depends on the individual instrument. Heavy rotor types perform best, but it is not necessary to try for speeds below 2000 rpm. (3) Some time is allowed for the rotor to become stable. No convection should be detectable in the cell(s). Thermal equilibrium should be reached before the measurements start. With the Spinco Model E, temperature control should be effected with the refrigeration unit only. (4) At an arbitrary time, which is wholly at the experimenter's discretion, the first concentration profile is recorded, which event marks the point $t = 0$ for that particular sample. This starting time is arbitrary for each sample, but once the measurements have begun, the schedule is rigid; the time intervals between successive scans (or photographs) must be equal for each sample, regardless of whether the samples are run simultaneously or consecutively. (5) At fixed time intervals, concentration profiles are recorded. Automatic ultraviolet scanning is ideal, especially in combination with digital data acquisition, although the older, less expensive, photographic ultraviolet system can be used if one is willing to invest more time and labor. Interference optics might be used if the concentration range to be studied lies outside the limits of workable optical densities; this method will also be found far more laborious than direct ultraviolet scanning with digital data recording. Useful time intervals depend on the diffusion coefficients involved, but usually lie between 8 and 32 min. The run(s) can be terminated at any time after enough data have been collected, but before the boundaries have smoothed out completely. (6) The radial distances of meniscus and cell bottom are measured for each sample. The values should be as nearly equal as possible. With the values for a and b obtained, eq 4 (*cf.* Theoretical section) is solved for p. Any iterative numerical procedure will do. In this laboratory, a computer program is used, which finds all roots of a general function; the program is based on the Muller method (Muller, 1956; Frank, 1958).[1] The smallest root found is used as the value of the parameter p. (7) All concentration profiles are sampled over the entire length of the liquid column, and at as many points as is practical (50 points is fairly minimal). If only graphical representations are available, either from direct scans, or in the form of densitometer tracings, the points must be read from these graphs. With direct digital output, the work involved is considerably less. (8) The composite Bessel function $B_0(pr)$ is calculated at each sample point, using the value of p which was determined earlier. The concentration at each sample point is multiplied by the radial distance, r, and the function value of $B_0(pr)$ at that point. A numerical integration is performed on these products, extending over the length of the cell.[1] These computations yield a single transform value for each sample, at each time interval. (9) The transform values for each sample are arranged in a row, so that a matrix is obtained with the number of rows being equal to the number of samples run, and with each column representing a specific time interval. Rank analysis is then performed upon this matrix,[1] and the numerical value found is equal to the number of species in the system.

[1] In this laboratory, three computer programs are used sequentially, for the determination of p, the transformation of the concentration profiles, and the rank determination. The programs are named ROBES, DCELERY, and NEWRANK, respectively, and can be obtained from the author upon request. The language is extended ALGOL for the Burroughs B5500 computer.

Discussion

The experimenter using the diffusion boundary transformation method should be aware of its limitations, as well as the liberties that may be taken.

As is the case with any experimental method, the results cannot be more accurate than the raw data. No species can, of course, be detected in concentrations which are too low to be significant in the measuring system. For example, a rank of two found for the matrix under investigation means that *at least* two species are present; it also means that *not more* than two species are present in concentrations above the limit imposed by the standard error of the measurement plus any additional errors compounded by the numerical processing of the data. As was pointed out before, error propagation throughout the rank analysis should be taken into consideration.

Automatic digital data acquisition, if available, is strongly recommended for both accuracy and convenience. Most commercially available data systems produce a magnetic or a punched paper tape, which can be processed directly by the computer.

Speed control at the low-speed operation of the centrifuge is unimportant, as long as the speed is sufficiently low. Rotor wobble must be prevented.

Good temperature control is mandatory; all samples belonging to the same set should be run at the same temperature. Even more important, though, is the absence of thermal fluctuations during the run. Changes in temperature generally cause convection, which upsets the diffusion process, and leads to erratic results. The restriction on temperature fluctuations applies here even more strictly than in sedimentation equilibrium, as the diffusion boundary transformation technique is a transport method.

The method depends on conservation of mass. It is, therefore, never permissible to let the solute form a pellet on the bottom of the cell; nor should leakage of the cell contents be permitted. Moreover, the concentration must be within the reading capability of the system at any point in the cell.

Ideality is one of the assumptions used in the development of the theory presented here, but actually the enumeration of species can still be performed in the same way for nonideal solutions. Of course, with the low concentrations which one normally uses with the ultraviolet scanner, nonideality is seldom of any importance.

The diffusion boundary transformation method is absolutely free of zero-time uncertainties. After a boundary has been formed, and the temperature has become constant, the moment of zero time may be chosen at the experimenter's convenience. The intervals are also arbitrary, but, once chosen for one cell, must be equal for all cells within a complete set. The intervals need not be constant throughout the run, however, so that a minor accident or mistake (which will happen), leading to the loss of one data scan, will not ruin the whole experiment. All that has to be done is eliminate the corresponding data scans from all other cells belonging to the set.

All flows have been expressed with reference to the volume-fixed frame, as is customary in all ultracentrifugation theory (and which is usually silently implied). As the practical measurements use the cell as the frame of reference, it is required that the partial specific volumes of all species present are constant throughout the experiment.

If ultraviolet scanning is used, Beer's law should be obeyed by the system, and the specific extinction coefficients of all species should be equal.

As the method depends entirely on diffusion of a boundary, it can be argued that a centrifuge is actually nonessential. Indeed, the method can be expected to give even more accurate results with a regular diffusion column, because of the greatly increased column length and the absence of all sources of error which are inherent in mechanical systems with moving parts. As the geometry of a regular diffusion column is rectangular, instead of sector shaped, a simple transform of the Fourier type could be used, where trigonometric functions would take the place of the Bessel functions. In a long column, no specific boundary conditions would have to be set up, *i.e.*, the column could be assumed to be of infinite length. These simplifications would lead to less elaborate arithmetic, and savings in computer time. The use of Fourier transforms for the evaluation of diffusion data on polydisperse, noninteracting systems, is well illustrated in the thorough and elaborate paper by Sundelöf (1965). Most laboratories are not set up for diffusion measurements: the necessary equipment for an efficient use of the transform method is rather specialized; a quartz diffusion column with direct ultraviolet scanning and digital output is required. This would present the experimenter with the problem of designing a scanning system with the photomultiplier and the light source moving along the column simultaneously. The method whereby the column itself moves, as described by Brumbaugh and Ackers (1968), works well with gel chromatography, but would be subject to the same possible vibration problems as the centrifuge, when used for free diffusion experiments. As analytical ultracentrifuges with ultraviolet scanner are quite widespread, it seems logical to use them for our purpose. A further reason for using a centrifuge is the possibility of extending the method to the case where sedimentation is not negligible. More recent theoretical developments have led to a new solution of the Lamm equation, in terms of the integral transform defined in eq 1. Preliminary analysis of the solution suggests that its application to interacting systems is feasible, and is likely to result in a more sensitive and versatile method. These developments, however, will be reported elsewhere.

Acknowledgments

Mr. Howard S. Warshaw's caustic persistence in pointing out the flaws in previous versions of the theory is greatly appreciated. Thanks are due to Dr. G. K. Ackers for valuable discussion. The highly efficient assistance of Miss Sophia C. De Bont with the computer analysis of the theoretical results is gratefully acknowledged.

References

Ackers, G. K. (1968), *J. Biol. Chem. 243*, 2056.
Ainsworth, S. (1961), *J. Phys. Chem. 65*, 1968.
Ainsworth, S., and Bingham, W. A. (1968), *Biochim. Biophys. Acta 160*, 10.
Brumbaugh, E. E., and Ackers, G. K. (1968), *J. Biol. Chem. 243*, 6315.
De Groot, S. R., and Mazur, P. (1962), Non-Equilibrium Thermodynamics, Amsterdam, North-Holland Publishing Co.
Frank, W. L. (1958), *J. Assoc. Computing Machinery 5*, 154.
Fujita, H. (1962), Mathematical Theory of Sedimentation Analysis, New York, N. Y., Academic Press.
Fujita, H., and Gosting, L. J. (1960), *J. Phys. Chem. 64*, 1256.
Godschalk, W. (1968), *J. Phys. Chem. 72*, 498.
Katchalsky, A., and Curran, P. F. (1965), Nonequilibrium Thermodynamics in Biophysics, Cambridge, Mass., Harvard Univ. Press.
Kegeles, G., and Sia, C. L. (1963), *Biochemistry 2*, 906.
Kharkevich, A. A. (1960), Spectra and Analysis, New York, N. Y., Consultants Bureau.
Markham, R. (1962), *Advan. Virus Res. 9*, 241.
Markham, R. (1967), *in* Methods of Virology, Vol. 2, Maramorosch, K., and Koprowski, H. Ed., New York, N. Y., Academic Press, p 275.
McMullen, D. W., Jaskunas, S. R., and Tinoco, I. (1967), *Biopolymers 5*, 589.
Muller, D. E. (1956), *Math. Tables Aids Comput. 10*, 208.
Sundelöf, L. O. (1965), *Ark. Kemi. 25*, 1.
Wallace, R. M. (1960), *J. Phys. Chem. 64*, 899.
Wallace, R. M., and Katz, S. M. (1964), *J. Phys. Chem. 68*, 3890.
Weber, G. (1961), *Nature (London) 190*, 27.
Winzor, D. J., and Scheraga, H. A. (1964), *J. Phys. Chem. 68*, 338.