Chapter 23
Analysis of Nonideal, Interacting, and Noninteracting Systems by Sedimentation Velocity Analytical Ultracentrifugation

Walter F. Stafford

Abstract Analysis of ideal associating systems has been described in detail previously (Rivas et al., Methods 19:194–212, 1999): the reader is referred to that article for the basic theory. We will extend that analysis to nonideal and self- and hetero-associating systems by adding terms for deviations from both thermodynamic ideality and hydrodynamic ideality. In this chapter we will consider several effects of non-ideality on the sedimentation process.

Keywords Analytical ultracentrifugation • Thermodynamic non-ideality • Hydrodynamic non-ideality • Self-association • Hetero-association • Diffusion • Sedimentation velocity • Nonlinear curve fitting

23.1 Background Theory

To put things in perspective, let's start with a quote from Williams et al. (1958):

> It is a well-known fact that the sedimentation methods have enjoyed a spectacular success in protein chemistry. It is now apparent that because of his enthusiasm for the transport method the protein chemist has on occasion allowed himself to be carried to some excesses. For instance, ideal equations descriptive of behavior in two-component systems with no volume change on mixing have been used to describe the experimental observations in multicomponent and not entirely ideal systems. Apparent single translational friction coefficients have been combined with other data and assumptions to provide information about the shape and volume of protein and polysaccharide molecules, when several such coefficients must have been involved.

This problem still exists some 50 odd years later. It is especially prevalent among the users of easy-to-use, "black-box" software packages that have become so popular. It is hoped that this chapter will provide a rigorous approach to the problems of treating non-ideality in associating systems. Some approximations are

W.F. Stafford (✉)
Boston Biomedical Research Institute, ℅12 Francis Avenue, Cambridge, MA 02138, USA
e-mail: wstafford3@walterstafford.com

S. Uchiyama et al. (eds.), *Analytical Ultracentrifugation*,
DOI 10.1007/978-4-431-55985-6_23

inevitably necessary to provide tractable relationships. These assumptions and their consequences must be kept in mind when analyzing data. Mike Johnson has outlined the "rules" that must be heeded when performing least squares analyses of any type of data (Johnson 1992), which are as follows:

1. The model must be correct.
2. The noise must be normally distributed with mean of zero.
3. The data must be free of systematic error.
4. The data set must have a large number of data points.
5. The observations must be independent.
6. The independent variable must be free of experimental uncertainty

It should be obvious that if the model is incorrect, the parameter values obtained from it will be meaningless. This will happen, for instance, if one were to try to fit to an interacting system with a noninteracting model or vice versa or a nonideal system with an ideal model.

23.1.1 Self-Associations

23.1.1.1 Self-Association: Two Species

A simple monomer to N-mer self-association can be described by two equations, a relation for mass action and a statement of conservation of mass for this one-component system:

$$\mathrm{nA} \rightleftharpoons \mathrm{A_n} \tag{23.1}$$

$$K_{1,n} = \frac{a_{A_n}}{a_A^n} = \left(\frac{\gamma_{A_n}}{\gamma_A^n}\right)\frac{C_{A_n}}{C_A^n} \tag{23.2}$$

$$C_A^o = C_A + nC_{A_n} \tag{23.3}$$

where a denotes the thermodynamic activity; γ_i, the molar activity coefficient; C, the molar concentration; K, the molar equilibrium constant; and C_A^o, the total concentration of component A.

23.1.1.2 Self-Associations: Multispecies

Higher-order self-associations or sequential associations can, in general, be represented by a series of reactions and one conservation relation since this is a one-component system:

$$2\mathrm{A_1} \rightleftharpoons \mathrm{A_2} \tag{23.4}$$

$$\mathrm{A_2}+\mathrm{A_1} \rightleftharpoons \mathrm{A_3} \tag{23.5}$$

$$A_3 + A_1 \rightleftharpoons A_4 \tag{23.6}$$

$$\ldots \tag{23.7}$$

$$A_{n-1} + A_1 \rightleftharpoons A_n \tag{23.8}$$

$$K_{i-1,n} = \left(\frac{\gamma_{A_n}}{\gamma_{A_{i-1}}\gamma_A}\right)\frac{C_{A_n}}{C_{A_{i-1}}C_A}; \qquad i = 2, 3, \ldots, n \tag{23.9}$$

$$C_A^o = C_{A_1} + 2C_{A_2} + 3C_{A_3} + 4C_{A_4} + \ldots + nC_{A_n} \tag{23.10}$$

23.1.2 Hetero-Associations

23.1.2.1 Hetero-Association: Bimolecular, Single Step

A simple two-component hetero-associating system can be represented by a mass action relation and two conservation of mass relations, one for each of the two components, A and B:

$$A + B \rightleftharpoons AB \tag{23.11}$$

$$K_{AB} = \left(\frac{\gamma_{AB}}{\gamma_A\gamma_B}\right)\frac{C_{AB}}{C_AC_B} \tag{23.12}$$

$$C_A^o = C_A + \frac{M_A}{M_{AB}}C_{AB} \tag{23.13}$$

$$C_B^o = C_B + \frac{M_B}{M_{AB}}C_{AB} \tag{23.14}$$

23.1.2.2 Hetero-Association: Bimolecular, Two Step

A somewhat more complicated, but frequently encountered, system is a two-step (e.g., antigen-antibody system) hetero-association of the following form:

$$A + B \rightleftharpoons AB \tag{23.15}$$

$$AB + B \rightleftharpoons AB_2 \tag{23.16}$$

$$K_{AB} = \left(\frac{\gamma_{AB}}{\gamma_A\gamma_B}\right)\frac{C_{AB}}{C_AC_B} \tag{23.17}$$

$$K_{AB_2} = \left(\frac{\gamma_{AB_2}}{\gamma_{AB}\gamma_B}\right)\frac{C_{AB_2}}{C_{AB}C_B} \tag{23.18}$$

$$C_A^o = C_A + \frac{M_A}{M_{AB}} C_{AB} + \frac{M_A}{M_{AB_2}} C_{AB_2} \tag{23.19}$$

$$C_B^o = C_B + \frac{M_B}{M_{AB}} C_{AB} + \frac{2M_B}{M_{AB_2}} C_{AB_2} \tag{23.20}$$

Other more complicated single- and multicomponent systems can be derived by simple extension of these relationships.

23.2 General Discussion of Non-ideality

Sedimentation velocity analysis of nonideal systems requires taking into account both hydrodynamic and thermodynamic non-ideality (Stafford and Sherwood 2004). Hydrodynamic non-ideality arises due to the displacement of solvent by the sedimenting macromolecules. Because the centrifuge cell is a closed system, the result is a "backflow" (a countercurrent, if you will) of solvent displaced by the macromolecule, which impedes its transport relative to the cell's coordinate system. The backflow is affected by the shape of the macromolecule and its charge and the ionic strength, which in turn determine its effective Stokes radius (Fuoss and Onsager 1961). This backflow leads to a decrease in both sedimentation coefficient and diffusion coefficient with increasing concentration. The backflow contribution is proportional to the concentration of macromolecule. It affects both sedimentation and diffusional transport equally (see Appendix 1). The reader is also referred to the monograph by Katchalsky and Curran, Chapter 9, especially Equations 9–19 (Katchalsky and Curran 1967) for further enlightenment.

23.2.1 Hydrodynamic Non-ideality

23.2.1.1 Single Macromolecular Component, Single-Species Systems

We can express the hydrodynamic non-ideality through the frictional coefficient. The frictional coefficient, f^o, at infinite dilution is given by

$$f^o = N_A 6\pi\eta_o R_s \tag{23.21}$$

where R_s is the Stokes radius, N_A is Avogadro's number, and η_o is the viscosity of water at 20 °C.

The concentration dependence of f is given by

$$f = f^o(1 + k_s c) \tag{23.22}$$

as a first-order function of concentration. From the definition of the sedimentation coefficient as

$$s = \frac{M(1 - v\rho)}{f} \tag{23.23}$$

we can express the concentration dependence of the sedimentation coefficient as

$$s(c) = \frac{s^o}{(1 + k_s c)} \tag{23.24}$$

where s^o is the value of s at infinite dilution, and in the absence of thermodynamic concentration dependence, we can likewise express the hydrodynamic concentration dependence of the diffusion coefficient as

$$D(c) = \frac{D^o}{(1 + k_s c)} \tag{23.25}$$

23.2.2 Thermodynamic Non-ideality

The chemical potential, μ_i, for species i, is given by

$$\mu_i = \mu_i^o + RTln(y_i c_i) \tag{23.26}$$

Taking the total differential of both sides and dividing through by RT, we have

$$\frac{d\mu_i}{RT} = dln(c_i) + dln(y_i) \tag{23.27}$$

Expanding in terms of the corresponding partial derivatives, we have an expression for the concentration dependence of the activity coefficient as a function of the concentration of all species present.

$$dln(y_i) = \sum_{j=1}^{n} \frac{\partial ln(y_i)}{\partial ln(c_j)} dln(c_j) \tag{23.28}$$

This can be rewritten as

$$dln(y_i) = \frac{\partial ln(y_i)}{\partial ln(c_i)} dln(c_i) + \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{\partial ln(y_i)}{\partial ln(c_j)} dln(c_j) \tag{23.29}$$

Now, substituting Eq. 23.29 into Eq. 23.27 , we have

$$d\mu_i/RT = dln(c_i) + \frac{\partial ln(y_i)}{\partial ln(c_i)} dln(c_i) + \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{\partial ln(y_i)}{\partial ln(c_j)} dln(c_j) \tag{23.30}$$

and rearranging

$$d\mu_i/RT = dln(c_i)\left(1 + \frac{\partial ln(y_i)}{\partial ln(c_i)}\right) + \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{\partial ln(y_i)}{\partial ln(c_j)} dln(c_j) \tag{23.31}$$

This equation is essentially Equation 9 of Goldberg (1953).

The driving force for diffusion is the gradient of the chemical potential, and so the diffusional flux, J_i^d, is proportional to the gradient of the chemical potential. From Fick's first law in terms of the chemical potential gradient, we can write:

$$J_i^d = -\frac{D_i c_i}{RT}\left(\frac{\partial \mu_i}{\partial r}\right)_t \tag{23.32}$$

Thermodynamic non-ideality, as we mentioned above, is manifest through the concentration dependence of the activity coefficient according to the following equations relating the gradient of the chemical potential to the concentration gradients. In cylindrical coordinates, in the case of diffusion in the absence of sedimentation, the continuity equation can be written as

$$\left(\frac{\partial c}{\partial t}\right)_\xi = -\nabla J_i^d \tag{23.33}$$

$$\nabla J_i^d = \frac{1}{r}\frac{\partial}{\partial r}\left[rJ_i^d\right] = \frac{\partial}{\partial \xi}\left[-r\frac{D_i c_i}{RT}\left(\frac{\partial \mu_i}{\partial r}\right)_t\right] \tag{23.34}$$

Expressing Eq. 23.33 in cylindrical coordinates and multiplying top and bottom of the right hand side by r, we have

$$\nabla J_i^d = \frac{\partial}{\partial \xi}\left[-2\xi\frac{D_i c_i}{RT}\frac{d\mu_i}{d\xi}\right] \tag{23.35}$$

where $\xi = r^2/2$.

Now, expanding the gradient of μ_i in terms of concentrations and activity coefficients, we have

$$\frac{1}{RT}\frac{d\mu_i}{d\xi} = \frac{dln(c_i)}{d\xi} + \sum_{j=1}^{n} \frac{\partial ln(y_i)}{\partial ln(c_j)}\frac{dln(c_j)}{d\xi} \tag{23.36}$$

This can be factored:

$$\frac{1}{RT}\frac{d\mu_i}{d\xi} = \frac{dln(c_i)}{d\xi} + \frac{\partial ln(y_i)}{\partial ln(c_i)}\frac{dln(c_i)}{d\xi} + \sum_{\substack{j=1 \\ j\neq i}}^{n} \frac{\partial ln(y_i)}{\partial ln(c_j)}\frac{dln(c_j)}{d\xi} \tag{23.37}$$

$$\frac{1}{RT}\frac{d\mu_i}{d\xi} = \frac{dln(c_i)}{d\xi}\left(1 + \frac{\partial ln(y_i)}{\partial ln(c_i)}\right) + \sum_{\substack{j=1 \\ j\neq i}}^{n} \frac{\partial ln(y_i)}{\partial ln(c_j)}\frac{dln(c_j)}{d\xi} \tag{23.38}$$

For a single species the concentration dependence of the activity coefficient has been treated historically by representing the concentration dependence with a polynomial in powers of concentration. This polynomial is often truncated after the first-order term, so that:

$$ln(y_i) = 2B_iM_ic_i + 0c^2 + \ldots \tag{23.39}$$

$$\frac{\partial ln(y_i)}{\partial c_i} = 2B_iM_i \tag{23.40}$$

substituting

$$\left(1 + c_i\frac{\partial ln(y_i)}{\partial c_i}\right) = (1 + 2B_iM_ic_i) \tag{23.41}$$

Both hydrodynamic non-ideality and thermodynamic non-ideality affect the diffusion coefficient, and we can write

$$D_i(c_i) = D_i^o\left[\frac{1 + c_i\frac{\partial ln(y_i)}{\partial c_i}}{1 + k_{s,i}c}\right] = D_i^o\left[\frac{1 + 2BMc_i}{1 + k_{s,i}c_i}\right] \tag{23.42}$$

However, in general, in a solution containing multiple species, we must include cross-terms to reflect the influence of all other species on each other. Expanding the non-ideality term in a first-order power series, we can write

$$ln(y_i) = 2\sum_{j=1}^{n} B_{i,j}M_jc_j \tag{23.43}$$

For example, for $n = 3, j = 1, 2, 3$, we have

$$ln(y_1) = 2B_{1,1}M_1c_1 + 2B_{1,2}M_2c_2 + 2B_{1,3}M_3c_3 \tag{23.44}$$

$$ln(y_2) = 2B_{2,1}M_1c_1 + 2B_{2,2}M_2c_2 + 2B_{2,3}M_3c_3 \tag{23.45}$$

$$ln(y_3) = 2B_{3,1}M_1c_1 + 2B_{3,2}M_2c_2 + 2B_{3,3}M_3c_3 \tag{23.46}$$

taking the total differential of each side:

$$dln(y_1) = 2B_{1,1}M_1dc_1 + 2B_{1,2}M_2dc_2 + 2B_{1,3}M_3dc_3 \tag{23.47}$$

$$dln(y_2) = 2B_{2,1}M_1dc_1 + 2B_{2,2}M_2dc_2 + 2B_{2,3}M_3dc_3 \tag{23.48}$$

$$dln(y_3) = 2B_{3,1}M_1dc_1 + 2B_{3,2}M_2dc_2 + 2B_{3,3}M_3dc_3 \tag{23.49}$$

So now we can write that

$$\frac{\partial ln(y_i)}{\partial ln(c_j)} = 2B_{i,j}M_jc_j; \tag{23.50}$$

giving nine, n^2, partial derivative terms relating the second virial self and cross coefficients, so that the so-called thermodynamic factor, let's call it N_i,

$$N_i = \left(1 + \sum_{j=1}^{n} \frac{\partial ln(y_i)}{\partial ln(c_j)}\right)$$

after substituting, becomes

$$N_i = \left(1 + \sum_{j=1}^{n} 2B_{i,j}M_jc_j\right) \tag{23.51}$$

Now, for three species:

$$N_1 = (1 + \{2B_{1,1}M_1c_1 + 2B_{1,2}M_2c_2 + 2B_{1,3}M_3c_3\}) \tag{23.52a}$$

$$N_2 = (1 + \{2B_{2,1}M_1c_1 + 2B_{2,2}M_2c_2 + 2B_{2,3}M_3c_3\}) \tag{23.52b}$$

$$N_3 = (1 + \{2B_{3,1}M_1c_1 + 2B_{3,2}M_2c_2 + 2B_{3,3}M_3c_3\}) \tag{23.52c}$$

in matrix form

$$\mathbf{N} = \left(1 + \mathbf{B}'\mathbf{c}\right) \tag{23.53}$$

and for 3 species, we have

$$\mathbf{N} = \begin{bmatrix} N_1 \\ N_2 \\ N_3 \end{bmatrix} \tag{23.54}$$

$$\mathbf{B}' = \begin{vmatrix} 1 & B'_{11} & B'_{12} & B'_{13} \\ 1 & B'_{21} & B'_{22} & B'_{23} \\ 1 & B'_{31} & B'_{32} & B'_{33} \end{vmatrix} \tag{23.55}$$

where $B'_{i,j} = 2B_{i,j}M_j$
and

$$\mathbf{c} = \begin{vmatrix} 1 \\ c_1 \\ c_2 \\ c_3 \end{vmatrix} \tag{23.56}$$

The full expressions for the diffusion coefficients become

$$D_1 = D_1^0 \frac{\left(1 + B'_{1,1}c_1 + B'_{1,2}c_2 + B'_{1,3}c_3\right)}{(1 + k_{1,1}c_1 + k_{1,2}c_2 + k_{1,3}c_3)} \tag{23.57a}$$

$$D_2 = D_2^0 \frac{\left(1 + B'_{2,1}c_1 + B'_{2,2}c_2 + B'_{2,3}c_3\right)]}{(1 + k_{2,1}c_1 + k_{2,2}c_2 + k_{2,3}c_3)} \tag{23.57b}$$

$$D_3 = D_3^0 \frac{\left(1 + B'_{3,1}c_1 + B'_{3,2}c_2 + B'_{3,3}c_3\right)}{(1 + k_{3,1}c_1 + k_{3,2}c_2 + k_{3,3}c_3)} \tag{23.57c}$$

It should be noted that in this treatment we have assumed that the cross diffusion coefficients, $D_{i,j}$, are sufficiently small that they can be ignored (see below).

Equations 23.57a, 23.57b, and 23.57c can be written in matrix notation as

$$\mathbf{D} = \mathbf{D^0} \frac{(1 + \mathbf{B'c})}{(1 + \mathbf{k_s c})} \tag{23.58}$$

where $\mathbf{B'}$ is called the "BM matrix" and $\mathbf{k_s}$ the "k_s" matrix

$$f_1 = f_1^o \left(1 + k_{1,1}c_1 + k_{1,2}c_2 + k_{1,3}c_3\right) \tag{23.59}$$

$$f_2 = f_2^o \left(1 + k_{2,1}c_1 + k_{2,2}c_2 + k_{2,3}c_3\right) \tag{23.60}$$

$$f_3 = f_3^o \left(1 + k_{3,1}c_1 + k_{3,2}c_2 + k_{3,3}c_3\right) \tag{23.61}$$

For the k_s matrix for hydrodynamic non-ideality, we have

$$\mathbf{f} = \mathbf{f^o} \left(1 + \mathbf{k_s c}\right) \tag{23.62}$$

Let

$$\mathbf{F} = (1 + \mathbf{k_s c}) \tag{23.63}$$

where (e.g., for 3 species)

$$\mathbf{F} = \begin{vmatrix} f_1/f_1^o \\ f_2/f_2^o \\ f_3/f_3^o \end{vmatrix} \tag{23.64}$$

$$\mathbf{k_s} = \begin{vmatrix} 1 & k_{s,11} & k_{s,12} & k_{s,13} \\ 1 & k_{s,21} & k_{s,22} & k_{s,23} \\ 1 & k_{s,31} & k_{s,32} & k_{s,33} \end{vmatrix} \tag{23.65}$$

$$\mathbf{c} = \begin{vmatrix} 1 \\ c_1 \\ c_2 \\ c_3 \end{vmatrix} \tag{23.66}$$

The $\mathbf{k_s}$ matrix and the $\mathbf{B}'$ matrices are implemented in SEDANAL as known, fixed parameters that must be obtained by independent measurements of the concentration interdependence of binary mixtures of the components in separate experiments. Experiments are underway to use these terms for the studies of proteins at high concentration (Correia 2015). These experiments are being carried out on model systems comprising binary mixtures of purified proteins in pairs initially to measure both the B' matrix and the Ks matrix with plans to apply them to proteins in serum.

23.2.3 Cross Diffusion Coefficients

Here we attempt to accommodate the cross diffusion coefficients $D_{i,j}$. Generally, the cross diffusion coefficients have been assumed to be insignificant and, therefore, mostly ignored. However, they may become important at higher concentrations found in biological fluids like serum. The cross diffusion coefficients might need to be included because the gradient of any one component in a mixture will become a driving force for diffusion of the other components in the mixture. First we consider the ideal case and then add in the thermodynamic non-ideality terms:

$$J_d^1 = -D_{11}\frac{c_1}{RT}\frac{\partial \mu_1}{\partial r} - D_{12}\frac{c_2}{RT}\frac{\partial \mu_2}{\partial r} - D_{13}\frac{c_3}{RT}\frac{\partial \mu_3}{\partial r} \tag{23.67a}$$

$$J_d^2 = -D_{21}\frac{c_1}{RT}\frac{\partial \mu_1}{\partial r} - D_{22}\frac{c_2}{RT}\frac{\partial \mu_2}{\partial r} - D_{23}\frac{c_3}{RT}\frac{\partial \mu_3}{\partial r} \tag{23.67b}$$

$$J_d^3 = -D_{31}\frac{c_1}{RT}\frac{\partial \mu_1}{\partial r} - D_{32}\frac{c_2}{RT}\frac{\partial \mu_2}{\partial r} - D_{33}\frac{c_3}{RT}\frac{\partial \mu_3}{\partial r} \tag{23.67c}$$

Ordinarily, we assume that

$$D_{i,j} = \delta_{i,j} D_{i,j} \tag{23.68}$$

where $\delta_{i,j}$ is the Kronecker delta.

Under nonideal conditions, taking into account the activity coefficients by expanding the chemical potential gradient in terms of concentrations and activity coefficients, we have

$$J_d^1 = -D_{11}\frac{c_1}{RT}\frac{\partial ln(y_1 c_1)}{\partial r} - D_{12}\frac{c_2}{RT}\frac{\partial ln(y_2 c_2)}{\partial r} - D_{13}\frac{c_3}{RT}\frac{\partial ln(y_3 c_3)}{\partial r} \tag{23.69a}$$

$$J_d^2 = -D_{21}\frac{c_1}{RT}\frac{\partial ln(y_1 c_1)}{\partial r} - D_{22}\frac{c_2}{RT}\frac{\partial ln(y_2 c_2)}{\partial r} - D_{23}\frac{c_3}{RT}\frac{\partial ln(y_3 c_3)}{\partial r} \tag{23.69b}$$

$$J_d^3 = -D_{31}\frac{c_1}{RT}\frac{\partial ln(y_1 c_1)}{\partial r} - D_{32}\frac{c_2}{RT}\frac{\partial ln(y_2 c_2)}{\partial r} - D_{33}\frac{c_3}{RT}\frac{\partial ln(y_3 c_3)}{\partial r} \tag{23.69c}$$

expanding

$$\begin{aligned} J_d^1 = &-D_{11}\frac{c_1}{RT}\frac{\partial ln(c_1)}{\partial r}\left(1 + \sum_{j=1}^{n}\frac{\partial ln(y_1)}{\partial ln(c_j)}\right) \\ &-D_{12}\frac{c_2}{RT}\frac{\partial ln(c_2)}{\partial r}\left(1 + \sum_{j=1}^{n}\frac{\partial ln(y_2)}{\partial ln(c_j)}\right) \\ &-D_{13}\frac{c_3}{RT}\frac{\partial ln(c_3)}{\partial r}\left(1 + \sum_{j=1}^{n}\frac{\partial ln(y_3)}{\partial ln(c_j)}\right) \end{aligned} \tag{23.70a}$$

$$\begin{aligned} J_d^2 = &-D_{21}\frac{c_1}{RT}\frac{\partial ln(c_1)}{\partial r}\left(1 + \sum_{j=1}^{n}\frac{\partial ln(y_1)}{\partial ln(c_j)}\right) \\ &-D_{22}\frac{c_2}{RT}\frac{\partial ln(c_2)}{\partial r}\left(1 + \sum_{j=1}^{n}\frac{\partial ln(y_2)}{\partial ln(c_j)}\right) \\ &-D_{23}\frac{c_3}{RT}\frac{\partial ln(c_3)}{\partial r}\left(1 + \sum_{j=1}^{n}\frac{\partial ln(y_3)}{\partial ln(c_j)}\right) \end{aligned} \tag{23.70b}$$

$$J_d^3 = -D_{31}\frac{c_1}{RT}\frac{\partial ln(c_1)}{\partial r}\left(1+\sum_{j=1}^{n}\frac{\partial ln(y_1)}{\partial ln(c_j)}\right)$$
$$-D_{32}\frac{c_2}{RT}\frac{\partial ln(c_2)}{\partial r}\left(1+\sum_{j=1}^{n}\frac{\partial ln(y_2)}{\partial ln(c_j)}\right) \quad (23.70c)$$
$$-D_{33}\frac{c_3}{RT}\frac{\partial ln(c_3)}{\partial r}\left(1+\sum_{j=1}^{n}\frac{\partial ln(y_3)}{\partial ln(c_j)}\right)$$

Now we can substitute the first-order approximate virial expansions:

$$\sum_{j=1}^{n}\frac{\partial ln(y_i)}{\partial ln(c_j)} = \sum_{j=1}^{n} 2B_{i,j}M_j c_j$$

Expanding as above (Eqs. 23.52a, 23.52b, and 23.52c) and substituting, we can write out explicitly for the flux of each component including both cross diffusion coefficients and cross second virial coefficients:

$$J_d^1 = -D_{11}\frac{c_1}{RT}\frac{\partial ln(c_1)}{\partial r}(1+2B_{1,1}M_1c_1+2B_{1,2}M_2c_2+2B_{1,3}M_3c_3)$$
$$-D_{12}\frac{c_2}{RT}\frac{\partial ln(c_2)}{\partial r}(1+2B_{2,1}M_1c_1+2B_{2,2}M_2c_2+2B_{2,3}M_3c_3) \quad (23.71a)$$
$$-D_{13}\frac{c_3}{RT}\frac{\partial ln(c_3)}{\partial r}(1+2B_{3,1}M_1c_1+2B_{3,2}M_2c_2+2B_{3,3}M_3c_3)$$

$$J_d^2 = -D_{21}\frac{c_1}{RT}\frac{\partial ln(c_1)}{\partial r}(1+2B_{1,1}M_1c_1+2B_{1,2}M_2c_2+2B_{1,3}M_3c_3)$$
$$-D_{22}\frac{c_2}{RT}\frac{\partial ln(c_2)}{\partial r}(1+2B_{2,1}M_1c_1+2B_{2,2}M_2c_2+2B_{2,3}M_3c_3) \quad (23.71b)$$
$$-D_{23}\frac{c_3}{RT}\frac{\partial ln(c_3)}{\partial r}(1+2B_{3,1}M_1c_1+2B_{3,2}M_2c_2+2B_{3,3}M_3c_3)$$

$$J_d^3 = -D_{31}\frac{c_1}{RT}\frac{\partial ln(c_1)}{\partial r}(1+2B_{1,1}M_1c_1+2B_{1,2}M_2c_2+2B_{1,3}M_3c_3)$$
$$-D_{32}\frac{c_2}{RT}\frac{\partial ln(c_2)}{\partial r}(1+2B_{2,1}M_1c_1+2B_{2,2}M_2c_2+2B_{2,3}M_3c_3) \quad (23.71c)$$
$$-D_{33}\frac{c_3}{RT}\frac{\partial ln(c_3)}{\partial r}(1+2B_{3,1}M_1c_1+2B_{3,2}M_2c_2+2B_{3,3}M_3c_3)$$

In summary, in general, for n components, we have a total of n^2 diffusion coefficients and n^2 first-order virial coefficients giving a total of $2n^2$ parameters to characterize diffusion in a nonideal n component system to first order in concentration. The total number of diffusion coefficients cannot be reduced by invoking the Onsager reciprocal relations (Onsager 1931a,b) which do not allow us to say that $D_{i,j} = D_{j,i}$.

Since, in general,

$$\begin{vmatrix} D_{11} & D_{12} & D_{13} \\ D_{21} & D_{22} & D_{23} \\ D_{31} & D_{32} & D_{33} \end{vmatrix} = \begin{vmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{vmatrix} \begin{vmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ \mu_{21} & \mu_{22} & \mu_{23} \\ \mu_{31} & \mu_{32} & \mu_{33} \end{vmatrix}$$

where the $L_{i,j}$ are the Onsager phenomenological coefficients, and

$$\mu_{i,j} \equiv \frac{\partial \mu_i}{\partial c_j}$$

For sedimentation velocity, we have Katchalsky and Curran (1967) (Eqns 9–29 through 9–50) the case that the observed sedimentation coefficients are functions of the chemical potential gradients of the other components leading to the following set of relations between the phenomenological coefficients, $L_{i,j}$, and the sedimentation coefficients. These equations result from a consideration of the gradient of the total potential gradient including both diffusion and sedimentation since for each component

$$\mu_i = \mu_i^o + \mu_i^c - M_i(1 - \rho\bar{v}_i)\omega^2 r^2/2 \tag{23.72}$$

where μ_i^c is the chemical potential, and the last term is the centrifugal potential. The standard reference potential term, μ_i^o, drops out upon differentiation, and we have

$$J_1 = s_1 c_1 \omega^2 r - D_{11}\frac{\partial c_1}{\partial r} - D_{12}\frac{\partial c_2}{\partial r} + D_{13}\frac{\partial c_3}{\partial r} \tag{23.73}$$

$$J_2 = s_2 c_2 \omega^2 r - D_{21}\frac{\partial c_1}{\partial r} - D_{22}\frac{\partial c_2}{\partial r} + D_{23}\frac{\partial c_3}{\partial r} \tag{23.74}$$

$$J_3 = s_3 c_3 \omega^2 r - D_{31}\frac{\partial c_1}{\partial r} - D_{32}\frac{\partial c_2}{\partial r} + D_{33}\frac{\partial c_3}{\partial r} \tag{23.75}$$

Leading to:

$$s_1 = \frac{1}{c_1}\left[L_{11}M_1(1 - \bar{v}_1\rho) + L_{12}M_2(1 - \bar{v}_2\rho) + L_{12}M_3(1 - \bar{v}_3\rho)\right]$$

$$s_2 = \frac{1}{c_2}\left[L_{21}M_1(1 - \bar{v}_1\rho) + L_{22}M_2(1 - \bar{v}_2\rho) + L_{22}M_3(1 - \bar{v}_3\rho)\right]$$

$$s_3 = \frac{1}{c_3}\left[L_{31}M_1(1 - \bar{v}_1\rho) + L_{32}M_2(1 - \bar{v}_2\rho) + L_{32}M_3(1 - \bar{v}_3\rho)\right]$$

where $L_{i,j}/c_i = D_{i,j}/RT$, and we have

$$s_1 = \left[\frac{D_{11}}{RT}M_1(1-\bar{v}_1\rho) + \frac{D_{12}}{RT}M_2(1-\bar{v}_2\rho) + \frac{D_{13}}{RT}M_3(1-\bar{v}_3\rho)\right]$$

$$s_2 = \left[\frac{D_{21}}{RT}M_1(1-\bar{v}_1\rho) + \frac{D_{22}}{RT}M_2(1-\bar{v}_2\rho) + \frac{D_{23}}{RT}M_3(1-\bar{v}_3\rho)\right]$$

$$s_3 = \left[\frac{D_{31}}{RT}M_1(1-\bar{v}_1\rho) + \frac{D_{32}}{RT}M_2(1-\bar{v}_2\rho) + \frac{D_{33}}{RT}M_3(1-\bar{v}_3\rho)\right]$$

Under conditions (e.g., high concentrations) in which the $L_{i,j}, i \neq j$, are not negligible, it will be necessary to include the $D_{i,j}$ in simulations using Lamm equation modeling. For cases in which the cross-terms are negligible, these equations reduce to the familiar Svedberg equations for each component.

23.2.4 Nonideal, Interacting Systems: Effects of Non-ideality on the Equilibrium Constant

We can write the equilibrium expression in terms of either molar or mass concentrations: in terms of activities and molar concentrations and then converting to weight concentrations we have, for a monomer-dimer system:
Molar:

$$K_{1,2} = \frac{a_2}{a_1^2} = \frac{\gamma_2 C_2}{(\gamma_1 C_1)^2} \tag{23.76}$$

Weight:

$$k_{1,2} = K_{1,2}\frac{M_2}{M_1^2} = \frac{y_2 c_2}{(y_1 C_1)^2} \tag{23.77}$$

$$k_{1,2,\text{obs}} = \frac{c_2}{c_2^2} = k_{1,2}\left(\frac{y_2}{y_1^2}\right)^{-1} \tag{23.78}$$

It is general practice, without much justification, to assume that

$$y_2 = y_1^2 \tag{23.79}$$

And in general for a multispecies self-associating system, it is mathematically convenient to assume

$$y_n = y_1^n \tag{23.80}$$

This approximation assumes that all species contribute the same non-ideality per unit mass (Tanford 1961) and obviously will not be valid if there is a significant change of overall charge or change in excluded volume upon association-disassociation. However, this approximation allows us to write, for the monomer-dimer case, that

$$k_{1,2,\text{obs}} = k_{1,2} \tag{23.81}$$

and for the general monomer to n-mer case:

$$k_{1,n,\text{obs}} = k_{1,n} \tag{23.82}$$

The reader is referred to the papers by Dennis Roark and David Yphantis (Roark and Yphantis 1969, 1971; Yphantis and Roark 1972) for a more thorough treatment of non-ideality in associating systems.

The situation is somewhat more complicated for hetero-associating systems since such an approximation generally cannot be easily made if the two proteins have significantly different charges and/or excluded volumes or if charges are canceled upon association. In the simplest bimolecular association, the three species will in general have sufficiently different properties (charge and/or excluded volume) that a simple relationship between activity coefficients usually cannot be written. One would need specific information concerning each species in the reaction. The necessary information could, in principle, be obtained by confined membrane electrophoresis (CME) (Filoti et al. 2015) for each of the species in the reaction for the effective charge on the macromolecule. Excluded volume differences would be more difficult to describe.

Nevertheless, some reasonable assumptions can be made in many cases. For example, in the two-component system

$$\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{AB}$$

assuming there are no significant changes in charge or excluded volume upon association, the non-ideality of species AB can be assumed to be the mass weighted average of the contributions from species A and B.

For mixtures of noninteracting components, the species (i.e., components) can be separated and studied independently to measure their individual properties and in pairs to study their mutual effects on each other to obtain the cross-terms. Measuring the cross-terms for an interacting system would be extremely difficult.

Sedimentation transport is described by the Lamm equation (Lamm 1929). For a single macromolecular species, i, we have for $c_i = c_i(r, t)$ that

$$\left(\frac{\partial c_i}{\partial t}\right)_r = -\frac{\partial}{r\partial r}\left[\omega^2 s_i(c_i)c_i - rD_i(c_i)\left(\frac{\partial c_i}{\partial r}\right)_t\right] \tag{23.83}$$

where c_i is the mass concentration of species i, r is the radius, t is the sedimentation time in seconds, ω is the angular velocity of the rotor in radians-sec^{-1}, s is the sedimentation coefficient in seconds, and D is the diffusion coefficient in centimeters^2-sec^{-1}

At this point it is useful to define some terms and point out the differences between a species and a component. A component, in the thermodynamic sense, is a chemical entity that can, in principle, be added to or removed from a solution independently of other species. A species is an individual chemical entity that may be a component or be a constituent of a component.

The Gibbs phase rule applied to the systems ordinarily encountered in the context of analytical ultracentrifugation tells us that, at constant temperature and pressure or in an incompressible solution, the number of macromolecular components is equal to the number of species minus the number of chemical reactions between them. For example, a monomer-dimer, self-associating system is a single-component system comprising two species. It's a single component because if one could remove the dimers, the remaining monomers would self-associate to form a mixture of monomer and dimers, and conversely, the dimers would dissociate to form a mixture of monomers and dimers. That is, monomers and dimers cannot be separated. At constant temperature and pressure, this system has one degree of freedom, namely, the total macromolecular concentration, which alone determines the composition (i.e., the fraction of monomers and dimers at equilibrium) given a particular value of the equilibrium constant.

For a two-component system, the number of degrees of freedom is 2. Again, at constant temperature and pressure, the composition of the solution is determined by the total concentration of each component given the equilibrium constants.

23.3 Curve Fitting: Numerical Solutions to the Lamm Equation

Curve fitting techniques allow us to combine data from several optical systems to fit to data from solutions that contain components that have different extinction properties. Model-independent methods, i.e., those not involving curve fitting, give us only the sum total of all the contributions from all species added together in one signal. For example, absorbance optics give us only the total absorbance for any particular solution of several species that may have quite different optical properties. Model-dependent methods, such as least squares curve fitting, allow us to extract the contributions from several species having different optical properties as long as we know the extinction coefficients of each species in a mixture.

To allow us to combine data sets from several optical systems, each of which may have different units and noise levels, we compute the sum of the weighted squares of the residuals, *WSSR*, over all the radial points, scans, and cells in a combined global fit (Stafford and Sherwood 2004), where the weighting factors are

the inverse of the variance of the data which, in general, is a function of radius. For example, the absorption optical system of the Beckman XL-A records the standard deviation of the absorbance or intensity along with its standard deviation at each radial position. The inverse of the square of the standard deviation can be used by SEDANAL to weight the squared residuals. With the interference optical system, the standard variation of the scan is essentially independent of position, and so a single weighting factor equal to the inverse of the standard deviation of the fringe displacement can be used. For fluorescence optics, both the magnitude and the standard deviation are much larger numerically than data from the other systems. However, when those data are normalized by weighting the squared residuals by the inverse of their variance, the signal-to-noise becomes comparable to weighted residuals from the other optical systems and they can be compared. Because the dimensionless weighted squares of the residuals are non-denominate numbers, they can be added together without violating any laws of mathematics.

$$WSSR = \frac{1}{LMN} \sum_{i=1}^{L} \sum_{j=1}^{M} \sum_{k=1}^{N} \left(\frac{dev_{i,j,k}^2}{\sigma_{i,j,k}^2} \right) \tag{23.84}$$

where N is the number of points over the range being fitted in each cell; M is the number of scans in each cell; L is the number of cells; $dev_{i,j,k}$ is the residual; and $\sigma_{i,j,k}$ is the standard deviation of the data. Thus, $(1/dev_{i,j,k})^2$, i.e., the inverse of the variance, becomes the weighting factor $w_{i,j,k}$.

$$WSSR = \frac{1}{LMN} \sum_{i=1}^{L} \sum_{j=1}^{M} \sum_{k=1}^{N} \left(w_{i,j,k} dev_{i,j,k}^2 \right) \tag{23.85}$$

For example, in SEDANAL we curve fit the Lamm equation to time difference curves, and so the residual is computed as

$$dev_{i,j,k} = \Delta S_{i,j,k}(r_i, t_j, t_{j+M/2}) - \sum_{l=1}^{l=n} \alpha_l \Delta C_{i,j,k,l}(r_i, t_j, t_{j+M/2}) \tag{23.86}$$

where $\Delta S_{i,j,k}(r_i, t_j, t_{j+M/2})$ is the time difference curve computed from the signal; n is the number of species in the model being fitted; α_l is the extinction coefficient for species l; and $\Delta C_{i,j,k,l}(r_i, t_j, t_{j+M/2})$ is the corresponding time difference curve computed from the current guesses in the solution of the Lamm equation. The reader is referred to the original paper for the details (Stafford and Sherwood 2004).

The time difference curves are computed as follows:

For the signal, S:

$$\Delta S_{i,j,k}(r_i, t_j, t_{j+M/2}) = S_{i,j,k}(r_i, t_{j+M/2}) - S_{i,j,k}(r_i, t_j) \tag{23.87}$$

And likewise, for the solutions, C, to the Lamm equation:

$$\sum_{l=1}^{l=n} \alpha_l \Delta C_{i,j,k,l}(r_i, t_j, t_{j+M/2}) = \sum_{l=1}^{l=n} \alpha_l C_{i,j,k,l}(r_i, t_{j+M/2}) - \sum_{l=1}^{l=n} \alpha_l C_{i,j,k,l}(r_i, t_j) \tag{23.88}$$

Weighting the residuals from different optical systems normalizes the residuals so that residuals from different optical systems can be combined in a global fit. Effectively, the result is that we minimize the reduced chi-square values over all the data sets. Since there are a large number of data points (usually several tens of thousands) and the noise on the data is normally distributed, this fitting procedure is the method of maximum likelihood (Bevington and Robinson 2003). This procedure maximizes the probability that the guesses for the parameters are correct.

Appendix 1: Hydrodynamic Non-ideality

It has been well established that at infinite dilution, the frictional coefficients for sedimentation and diffusion are equal (Schachman 1959). Above we claim that their dependence on concentration is also the same. We can show this to be true with a proof by contradiction, namely, that the hydrodynamic concentration dependencies of sedimentation and diffusion, k_{sed} and k_{diff}, are equal. We start by assuming they are not equal (i.e., $k_{\text{sed}} \neq k_{\text{diff}}$) and proceed to show that this assumption leads to a contradiction. This hydrodynamic concentration dependence arises purely because of the frame of reference we are using, i.e., the cell is of constant volume, and so displaced solvent has to move against the macromolecule and into the volume that was occupied by the macromolecule while the macromolecule is translating in response to the forces acting upon it, whether or not the force arises from the gradient of the centrifugal potential (the driving force for sedimentation) or the gradient of the chemical potential (the driving force for diffusional transport) (Katchalsky and Curran 1967). This relationship has been postulated in the past (Harding and Johnson 1985) but never explicitly proven.

Consider the fluxes arising from sedimentation and diffusion:

$$J_{\text{sed}} = \omega^2 rc \left[\frac{s_o}{1 + k_{\text{sed}} c} \right] \tag{23.89}$$

$$J_{\text{diff}} = -D_o \left[\frac{\left(1 + \frac{\partial ln(y)}{\partial ln(c)}\right)}{1 + k_{1\text{diff}} c} \right] \left(\frac{\partial c}{\partial r} \right)_t \tag{23.90}$$

At sedimentation equilibrium, the flux due to sedimentation is equal to the flux due to diffusion throughout the cell. And we have that

$$J_{\text{sed}} = J_{\text{diff}} \tag{23.91}$$

and we have

$$\omega^2 rc\left[\frac{s_o}{1+k_{\text{sed}}c}\right] = -D_o\left[\frac{\left(1+\frac{\partial ln(y)}{\partial ln(c)}\right)}{1+k_{\text{diff}}c}\right]\left(\frac{\partial c}{\partial r}\right)_t \tag{23.92}$$

This can be seen from the Lamm equation (Eq. 23.83) by setting $(\partial c/\partial t)_r = 0$ and rearranging:

$$\omega^2\frac{s_o}{D_o} = \left[\frac{1+k_{\text{sed}}c}{1+k_{\text{diff}}c}\right]\frac{1}{rc}\left(\frac{\partial c}{\partial r}\right)_t\left(1+\frac{\partial ln(y)}{\partial ln(c)}\right) \tag{23.93}$$

This equation agrees with the standard thermodynamic derivation if and only if k_{sed} is equal to k_{diff}.

$$\omega^2\frac{s_o}{D_o} = \frac{1}{rc}\left(\frac{\partial c}{\partial r}\right)_t\left(1+\frac{\partial ln(y)}{\partial ln(c)}\right) \tag{23.94}$$

Invoking the Svedberg equation, we have

$$\frac{M(1-\bar{v}\rho)\omega^2}{RT} = \frac{1}{rc}\left(\frac{\partial c}{\partial r}\right)_t\left(1+\frac{\partial ln(y)}{\partial ln(c)}\right) \tag{23.95}$$

which agrees with the thermodynamic derivation given by Williams et al. (1958). And so, it must be the case that

$$\left[\frac{1+k_{\text{sed}}c}{1+k_{\text{diff}}c}\right] = 1$$

Therefore, our initial assumption that $k_{\text{sed}} \neq k_{\text{diff}}$ must have been false, and we have proven that k_{sed} is equal to k_{diff} and that any difference in the concentration dependence of sedimentation and diffusion arises only through the thermodynamic non-ideality term which in turn arises from the concentration dependence of the activity coefficients (cf. Eq. 23.41) of each of the species present in the solution (Goldberg 1953; Sherwood and Stafford 2016). The equality of k_{sed} and k_{diff} has been demonstrated experimentally for sucrose (LaBar and Baldwin 1963). It should be pointed out that $k_{\text{sed}}c$ and $k_{\text{diff}}c$ (i.e., linear dependence on concentration) can be replaced by any function of concentration, $g(c)$, that describes the concentration dependence of the observed frictional coefficient.

References

Bevington PR, Robinson DK (2003) Data reduction and error analysis for the physical sciences, 3rd edn. McGraw-Hill, New York

Correia JJ (2015) An examination of the Johnston-Ogston effect using the AVIV-FDS and SEDANAL global direct boundary fitting. In: Uchiyama S et al. (eds) Analytical ultracentrifugation: instrumentation, software, and applications. Springer, Tokyo

Filoti DI, Shire SJ, Yadav S, Laue TM (2015) Comparative study of analytical techniques for determining protein charge. J Pharm Sci 104:2123–2131

Fuoss RM, Onsager L (1961) Thermodynamic potentials of symmetrical electrolytes. PNAS 47:818–825

Goldberg RJ (1953) Sedimentation in the ultracentrifuge. J Phys Chem 57:194–197

Harding SE, Johnson P (1985) The concentration-dependence of macromolecular parameters. Biochem J 231:543–547

Johnson ML (1992) Why, when, and how biochemists should use least squares. Anal Biochem 206(2):215–225. Review

Katchalsky A, Curran PF (1967) Nonequilibrium thermodynamics in biophysics. Harvard University Press, Cambridge

LaBar FE, Baldwin RL (1963) The sedimentation coefficient of sucrose. J Am Chem Soc 85(20):3105–3108

Lamm O (1929) Die differentialgleichung der ultrazentrifugierung. Ark Mat Astron Fys 21B(2): 1–4

Onsager L (1931a) Reciprocal relations in irreversible processes i. Phys Rev 37:405–426

Onsager L (1931b) Reciprocal relations in irreversible processes. ii. Phys Rev 38:2265–2279. doi:10.1103/PhysRev.38.2265. http://link.aps.org/doi/10.1103/PhysRev.38.2265

Rivas G, Stafford WF, Minton A (1999) Characterization of heterologous protein-protein interactions using analytical ultracentrifugation. Methods 19:194–212

Roark DE, Yphantis DA (1969) Studies of self-associating systems by equilibrium ultracentrifugation. Ann NY Acad Sci 164(1):245–278

Roark DE, Yphantis DA (1971) Equilibrium centrifugation of nonideal systems. The Donnan effect in self-associating systems. Biochemistry 10(17):3241–3249

Schachman HK (1959) Ultracentrifugation in biochemistry. Academic, New York

Sherwood PJ, Stafford WF (2016) Sedanal: model-dependent and model independent analysis of sedimentation data. In: Uchiyama S, Arisaka F, Stafford WF, Laue T (eds) Analytical ultracentrifugation: instrumentation, software, and applications. Springer, Tokyo, pp xx–xx

Stafford WF, Sherwood PJ (2004) Analysis of heterologous interacting systems by sedimentation velocity: curve fitting algorithms for estimation of sedimentation coefficients, equilibrium and kinetic constants. Bioph Chem 108:231–243

Tanford C (1961) Physical chemistry of macromolecules. Wiley, New York

Williams JW, Van Holde KE, Baldwin RL, Fujita H (1958) Theory of sedimentation analysis. Chem Rev 58:715–744

Yphantis DA, Roark DE (1972) Equilibrium centrifugation of nonideal systems. Molecular weight moments for removing the effects of nonideality. Biochemistry 11(16):2925–2934