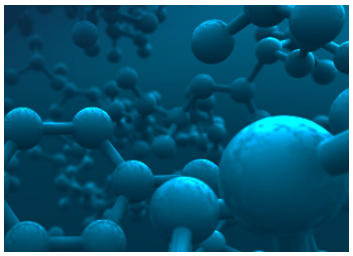
ASTRA 7
User's Guide



Version 7.1
(M1006 Rev. D)

WYATT TECHNOLOGY

6330 Hollister Avenue
Santa Barbara, CA 93117
Tel: +1 (805) 681-9009
Fax: +1 (805) 681-0123
Web: www.wyatt.com

E Light Scattering Theory

This appendix reviews the basic theory of light scattering and how the ASTRA software determines the molar masses and root mean square radii of a sample whose light scattering properties have been measured. The text covers basic quantities, the relation to measurements, calibration and normalization, and determination of molar masses, sizes and distributions.

CONTENTS **PAGE**

Introduction 377
Basic Quantities 378
Measured Quantities and Calibration 379
Normalization 383
Concentration Calculation Methods 386
Determination of Molar Mass and Sizes 389
Molar Mass and rms Radius Moments 395
Uncertainties in Calculated Quantities 398
Out of Range Values 399
Differential Distribution Calculations 399
Branching Calculations 400

Introduction

Perhaps the most important application of the ASTRA program is its ability to convert measurements of a fractionated sample, passing through appropriate instrumentation, into an accurate representation of the molar masses and sizes present in the sample.

Although size-exclusion chromatography (SEC) provides good separation of molecules based on their hydrodynamic radii, before the advent of light scattering techniques it had been necessary to calibrate SEC columns using standard samples of known molar mass in order to determine only the molar masses present in an unknown sample. Unfortunately, appropriate standards having the same composition and effective conformations as the unknown specimen have never really been available.

On the other hand, if the value of the differential refractive index increment (dn/dc) or the total mass of eluting solute is known, light scattering measurements can provide an absolute measurement of molar mass when used in series with a concentration-sensitive detector such as a refractive index (RI) detector.

Thus, light scattering measurements effectively provide a column "calibration curve" for every sample, obviating time-consuming, conformation-dependent calibration procedures. When techniques such as reverse phase chromatography are used, separation is not based on molecular hydrodynamic size, and calibration techniques based on known standards are useless.

Basic Quantities

The light scattered into a detector centered at angle θ per unit solid angle subtended by said detector is called the Rayleigh ratio or R_θ, and is defined as:

$$R_\theta = \frac{I_\theta r^2}{I_0 V} \qquad (1)$$

where I_θ is the scattered intensity, I_0 is the intensity of the incident beam, V is the illuminated volume of the scattering medium from which the detector at θ collects light, and r is the distance between V and the detector.

Eq. (1) implies that the dimension of R_θ is (length)$^{-1}$.

Note: It is assumed in Eq. (1) and throughout this manual that the incident light is always vertically polarized; that is, its electric field is perpendicular to the plane in which the angular variation of the light scattered from the sample is measured. All DAWN instruments are supplied with such vertically polarized laser sources.

More generally, R_θ usually refers to the *excess* scattering from a solution above that scattered by the solvent alone as follows:

$$R_\theta = \frac{\left(I_\theta - I_{\theta,solvent}\right) r^2}{I_0 V} \qquad (2)$$

where $I_{\theta,solvent}$ is the scattered intensity into angle θ from the solvent.

Thus scattering from a pure solvent is described by Eq. (1), while the scattering in excess above the solvent is described by Eq. (2).

Eq. (2) effectively describes the scattering after subtracting the "baseline" corresponding to the pure solvent.

As we shall see in the sections that follow, measurement of R_θ at a number of different angles combined with the corresponding molecular concentration provides the data by which the weight average molar mass and size of the solute molecules are determined.

Measured Quantities and Calibration

How do we measure R_θ? Photodetectors used in static light scattering instruments generate electrical voltages or currents that are proportional to the intensity of incident light. In order to relate the electrical signal to Rayleigh ratio, we must calibrate our detectors. This calibration is not difficult and will be explained shortly.

Furthermore, in the DAWN and miniDAWN instruments, each detector "sees" a slightly different scattering volume V and subtends a different solid angle with respect to the scattering volume. Refractive index differences among various solvents and sample cells also contribute to differences in scattering volumes.

We correct for the slightly different scattering volumes and subtended angles observed by each detector by using the proportionality between R_θ and I_θ in Eq. (1) to derive a calibration factor which gives the correct value of R_θ for a known scatterer. Several common solvents have been thoroughly studied, and their Rayleigh ratios are well known, allowing us to use a pure solvent as the calibration standard. Using pure solvent as the scattering standard makes the calibration completely independent of any particular sample.

The simplest and most robust calibration procedure utilizes scattering at 90° to the incident light beam. We combine the detector sensitivity with the geometric volume and solid angle factors into a single Configuration Specific Calibration Constant, called A_{CSCC}.[1] By using the proportionality between detector voltage and light intensity, Eq. (1) may be expressed as:

$$R_{90^\circ} = A_{CSCC}\left(\frac{V_{90^\circ} - V_{90^\circ dark}}{V_{laser} - V_{laser,dark}}\right) \qquad (3)$$

where V_{90° and $V_{90^\circ dark}$ are the 90° detector signal voltage and its dark offset voltage, respectively. V_{laser} and $V_{laser,dark}$ are the laser monitor signal and its dark offset, respectively. Dark offsets are obtained by turning off the laser and measuring the detector signal. Division by the laser monitor signal compensates for any changes in laser intensity due to power supply fluctuations, temperature drift, laser aging, etc.

As an example, suppose we calibrate with toluene. Pure filtered toluene has a Rayleigh ratio of 1.406 x 10^{-5} cm^{-1} at a wavelength of 632.8 nm.[2] Suppose that using our DAWN or miniDAWN sample cell we observe a 90° scattering signal of about 1 V. The laser monitor signal is factory-set to be

1. The symbol A_{CSCC} is not displayed in the software.
2. W. Kaye and J.B. McDaniel, "Low-angle laser light scattering—Rayleigh factors and depolarization ratios," *Applied Optics*, vol. 13, No. 8, 1974, pp. 1934–1937.

near 5 V, and the dark offsets are much smaller than this, so Eq. (3) implies $A_{CSCC} \approx 7.0 \times 10^{-5}$ cm^{-1}. Of course, this is just an example, and the measured constant may be quite different depending on the instrument and conditions (laser wavelength, etc.).

Toluene provides a large scattering signal; in fact, toluene has the highest Rayleigh ratio of any of the common solvents, and is thus highly desirable for use as a calibrator. Many other solvents can, theoretically, be used for calibration of the DAWN or miniDAWN but we do not recommend them.

The astute reader will point out that since we know the Rayleigh ratio for toluene, and since the scattering from toluene is relatively large, we ought to be able to calibrate with toluene, measure our samples in water, and still obtain correct results. Unfortunately this simplistic procedure ignores the geometric factors affecting the volume of scattering molecules seen by the 90° detector, as well as the solid angle it subtends and reflections at interfaces between different materials such as air-glass and solvent-glass. These factors depend on the refractive index of both the solvent and the glass of which the sample cell is made (see the DAWN or miniDAWN Hardware Manual).

Thus the "constant" A_{CSCC} is really dependent on the solvent type and cell type. That is why we call it a configuration specific calibration constant. To allow users to calibrate with one solvent and/or cell and make measurements with another, we must have an "instrument" constant that is truly independent of these changing factors and is instead only a function of the particular instrument and the sample cell geometry.

This instrument constant A_{inst} is related to A_{CSCC} as follows:

$$A_{\mathrm{CSCC}} = A_{\mathrm{inst}}\left(\text{Reflection correction}\right)\left(\text{Geometry correction}\right) \qquad (4)$$

The reflection correction represents the reflective losses at each interface in the sample cell. For example, the incident laser beam loses intensity at the air-glass interface of the sample cell, and the glass-solvent interface as well. Similarly, the scattered light that is to be detected also suffers from reflective losses at the solvent-glass and glass-air interfaces as it leaves the sample cell. If the solvent is changed, or a different cell is used, these reflective losses will change, and hence it is necessary to correct for them if A_{inst} is to be independent of solvent and cell glass.

The reflection correction is calculated from changes in the transmitted intensity between media 1 and 2 with indices of refraction n_1 and n_2, respectively. The transmitted intensity from medium 1 into medium 2 is given by the Fresnel equations as:

$$T_{12} = \frac{4n_1 n_2}{\left(n_1 + n_2\right)^2} \qquad (5)$$

If g represents the flow cell glass, s represents the solvent, and a represents air, then the complete reflection correction can be written as:

$$\text{Reflection correction} = \frac{1}{T_{sg}^2 T_{ga}^{N^*}} \tag{6}$$

where N^* is the number of uncoated glass-air interfaces the incident and scattered light traverse. It is assumed that the reflective losses at an anti-reflection coated interface (such as the air-glass laser window interface) are negligible.

The "Geometry correction" for a sample cell is not as easily determined as the reflection correction. There are examples of analytical expressions derived for simple cell geometries[1,2], but there are no exact analytical expressions for practical cells. In addition, these analytical expressions are valid for conditions in which the source of scattered light is either a point source or a completely illuminated volume, neither of which hold for the current scattered light source—a collimated laser beam. Therefore, the geometry correction has been calculated for the DAWN flow cell using the paraxial approximation, and for scintillation vials and cuvettes using computer ray-tracing simulations based upon the exact geometry of the sample cell, laser beam, and detection optics for the 90 degree detector in the DAWN and miniDAWN instruments.

For the DAWN flow cell, the resulting geometry correction goes as $n_s n_g$, that is, the index of refraction of the solvent times the index of refraction of the glass, respectively. The complete expression taking into account the reflection and geometry corrections is therefore:

$$A_{\text{CSCC}} = A_{\text{inst}} \frac{n_s n_g}{T_{sg}^2 T_{ga}} \tag{7}$$

for the standard flow cell with an antireflection coated laser entrance and exit window and uncoated exit surface to the detector, and the transmission terms are calculated using Eq. (5).

When using a scintillation vial, both the geometry and solvent-glass reflection corrections were folded into the ray tracing calculations, so only the factor for the reflection correction due to the two uncoated glass-air interfaces of the scintillation vial are in the final expression. The resulting complete formula for the scintillation vial is:

$$A_{\text{CSCC}} = A_{\text{inst}} \frac{n_s^{1.797}}{T_{ga}^2} \tag{8}$$

1. C.I. Carr, Jr. and B.H. Zimm, "Absolute Intensity of Light Scattering from Pure Liquids and Solutions", *J. Chem. Phys.,* vol. 18, pp. 1616-1626 (1950).
2. J.J. Hermans and S. Levinson, "Some Geometric Factors in Light-Scattering Apparatus", *J. Opt. Soc. Am.*, vol. 41, pp. 460-464 (1951).

For the MicroCuvette, both the geometry and solvent-glass reflection corrections were folded into the ray tracing calculations. Both the entrance and exit windows for the MicroCuvette are anti-reflection coated, so there are no explicit reflection correction terms in the final equation. The resulting complete formula for the MicroCuvette is

$$A_{\mathrm{CSCC}} = A_{\mathrm{inst}} n_s^{1.983} \qquad (9)$$

During the performance of a calibration, ASTRA calculates a configuration-specific constant from Eq. (3), but this number is never seen. It is immediately converted to the instrument constant A_{inst} via Eq. (7), Eq. (8), or Eq. (9) depending on the cell type. The A_{inst} value is reported as the "Calibration Constant" and is the value entered in the DAWN or miniDAWN profile.

If at some later time A_{inst} is changed in the instrument profile, ASTRA will recalculate A_{CSCC}. ASTRA also recalculates A_{CSCC} if the solvent or cell type is changed. This process readily enables one to calibrate with one solvent and make measurements with another, while the software efficiently handles all the details.

Changing the sample cell requires recalibration of the MALS detector with the specific sample cell intended to measure the sample of interest.

The calibration measurements should be made with *great care* as the accuracy of all other measurements depends upon them. As long as the system is left undisturbed it is not necessary to recalibrate, but we advise making occasional checks using a standard reference molecular species, as photodiode sensitivity may change with age. The calibration should be performed with HPLC-grade toluene filtered through the smallest available filter (0.02 µm) immediately before making the measurement using the ASTRA program. The cleanliness of the cell is vital for this purpose. Be sure to leave the DAWN or miniDAWN instrument and the laser switched on for one hour before making any measurements.

Normalization

After performing the calibration procedure in toluene, we have calibrated the 90° detector in an absolute sense: the calibration is totally independent of any sample we might wish to study. In other words, we can measure R_θ accurately for any solvent or sample, assuming it gives a large enough signal. Furthermore, the calibration can be traced directly to the scattering from pure, well-understood solvents.

So far we have ignored all angles except 90°. Each detector has its own geometric factors and angular sensitivity to measured light intensity. Furthermore, these effects vary from solvent (and sample) to solvent. We would like to quantify this effect so that we can correct for it. If not, we will mistake solvent and geometric readhead effects for characteristics of our sample, resulting in erroneous molar mass and size.

Therefore, we use a set of *normalization coefficients* N_θ to relate each detector to the 90° detector. These coefficients must be determined using the *same temperature* and the *same solvent* that are used for the actual sample measurement, since the refractive index of the solvent changes the scattering angles and the geometric factors for each detector.

For purposes of normalization, we must employ a sample that is an isotropic scatterer (one which scatters equally in all directions), so that we can be sure that the variations measured are due to detector geometry and not some interaction of the sample with the light. Particles whose size is much smaller than the wavelength of the vertically polarized light incident upon them are often called Rayleigh particles and scatter such incident light isotropically. The normalization coefficient for the 90° detector is assigned a value of 1.0, while the other detectors are adjusted by varying amounts to yield identical values of R_θ at all angles.

The process of normalization is quite simple. We assume that the 90° detector has already been calibrated as described above. To normalize, we introduce an isotropic solute (*i.e.*, producing an R_θ is independent of θ) and compute a set of coefficients so that each detector gives the same R_θ as the 90° detector when its signal is multiplied by its correct normalization coefficient. Expressed algebraically we have:

$$R_\theta = N_\theta A_{CSCC}\left(\frac{V_\theta - V_{\theta,\,dark}}{V_{laser} - V_{laser,\,dark}}\right) \qquad (10)$$

For Eq. (3) and Eq. (10) to agree when $\theta = 90°$, N_{90} must be exactly unity.

Thus Eq. (10) gives us a way to calculate Rayleigh ratios at any detector angle. We recommend normalizing with a low molar mass sample whose constituents all have radii less than about 5 nm. (Molecules this small scatter nearly isotropically as discussed previously.) For organic solvents, small polystyrene samples are generally used with molar masses less than

about 30,000 g/mol. For aqueous solvents and buffers, dextran with a weight average molar mass of 10,000 g/mol or bovine serum albumin (BSA) with a molar mass of 66,400 g/mol may be used.

In practice we need not measure the various detector dark offsets $V_{90,dark}$ of Eq. (10). This is because the instrument is typically used to study samples in solution, not solvents by themselves. Thus we are interested in the *excess* Rayleigh ratio of the eluting sample, compared with the baseline of solvent alone. We therefore use an alternative form of Eq. (10):

$$R_\theta = N_\theta A_{CSCC}\left(\frac{V_\theta - V_{\theta,\,baseline}}{V_{laser} - V_{laser,\,dark}}\right) \qquad (11)$$

where R_θ is the excess Rayleigh ratio, and $V_{\theta,baseline}$ is the detector voltage at a great distance from any solute peak. The quantity $V_{\theta,baseline}$ is the scattering signal from the solvent alone [*cf.* Eq. (2)] plus the detector dark offset. Eq. (11) is used by ASTRA.

Implementation of Normalization

ASTRA provides two normalization techniques: “Area” normalization is the default used in ASTRA 6 and ASTRA 7. "Standard" normalization was used in older software versions of ASTRA.

Area Normalization

When using Area Normalization, the results are calculated from the integration of the Rayleigh Ratio peak as follows, rather than just using the peak apex as in Standard Normalization:

1. Select a sample peak to use for normalization.
2. Integrate the Rayleigh Ratios over the entire peak.
3. Repeat the integration for each light scattering detector.
4. Any negative results are set to 1.0.
5. Finally, set the normalization coefficients for all detectors equal to the result of the integration for the current detector divided by the result for the 90° detector. This forces the 90° detector to equal 1.0.

Standard Normalization

In practice, $V_\theta - V_{\theta,baseline}$ is not determined from a single data slice, but from the result of the following steps:

1. Select a sample peak to use for normalization.
2. Using the collected data points for the center half of the peak (that is, the half of the peak centered on the peak apex), fit them to a 6th order polynomial of the form.

$$y = a_0x + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5 + a_6x^6 \qquad (12)$$

Note that there must be at least seven points in the “center half” of the selected peak for the normalization calculation to run.

3. The apex of the fit curve, y in Eq. (12), provides $V_\theta — V_{\theta,baseline}$. The maximum y is found iteratively by using the x value for each slice used in the fit in Eq. (12), and selecting the largest resulting y. This method is used due to the relatively small number of points typically involved, and to preclude the chance of encountering another local maxima.
4. Repeat steps 2 and 3 first for the 90° degree detector, then for each light scattering detector for which a normalization coefficient is to be calculated.
5. Set any negative y values to 1.0.
6. Finally, divide the results of Eq. (12) for all detectors by the result of Eq. (12) for the 90° detector. This yields the desired normalization coefficients.

Concentration Calculation Methods

The concentration of flowing samples may be determined via any online concentration detector, such as UV/Vis absorption, differential refractive index, or fluorescence detectors. In batch methods, where concentration detectors cannot be used, the concentration values must be entered manually into the appropriate fields.

Optilab rEX, T-rEX, and UT-rEX

ASTRA communicates digitally with the Optilab to obtain the difference in the refractive index of the sample and reference cells. Since the reference cell usually contains pure solvent, this difference reflects the change in refractive index (dRI) due to the solute, which can be related to its concentration via dn/dc, the refractive index increment of the specific solute in the specific solvent. The value of dn/dc may be found in the literature for many solutes and solvents, calculated from a mass average of dn/dc values of components of the solute, or measured empirically with an Optilab.

Even if the Optilab's reference cell does not contain pure solvent, the solute concentration is determined by subtracting the dRI measured in pure solvent from the dRI measured in the presence of solute, given by Eq. (13):

$$\Delta c = \frac{(d_{solute} - d_{solvent})}{dn/dc} \qquad (13)$$

Other Concentration Detectors

Using any other concentration detector requires connecting an analog output signal, which is proportional to the signal measured by the concentration detector, to an analog input (AUX) channel on one of the Wyatt detectors used in the measurement, such as the DAWN, miniDAWN, Optilab, or ViscoStar. A concentration calibration factor alpha, which relates the analog voltage difference on the AUX channel to concentration, must be determined.

There are several distinct options for configuring an analog concentration signal, such as UV and generic RI concentration detectors. These options are described in the following sections.

Determining Concentration by Generic RI

A generic RI concentration detector can be used with either a known dn/dc value and a known AUX calibration constant or with only a known AUX calibration constant if you assume that 100% mass recovery occurs.

Known dn/dc and Known AUX Calibration Constant

The sample concentration for each data slice is determined by the RI concentration detector assuming a constant dn/dc value across the sample peak.

This is the default method, and is the one we recommend. It requires known values for the "RI" calibration constant α and the differential refractive index increment dn/dc (in mL/g). It does not require that the total injected mass be known and is independent of an accurate flow rate. The quantity α can be determined by injecting a sample with known dn/dc into the refractometer at a few different concentrations, and the dn/dc value may be found in the literature or measured using an Optilab instrument.

For the i^{th} slice, the change in refractive index compared to pure solvent is given by:

$$\Delta n_i = \alpha(V_i - V_{i,baseline}) \tag{14}$$

where V_i and $V_{i,baseline}$ are the RI signal and baseline voltages, respectively. Dividing Δn_i by dn/dc gives the change in concentration of solute, compared to the baseline, for that slice:

$$\Delta c_i = \frac{\Delta n_i}{dn/dc} = \frac{\alpha(V_i - V_{i,baseline})}{dn/dc} \tag{15}$$

Since the baseline represents the signal from the pure solvent, $\Delta c_i = c_i$. Once the concentration is known, the mass w_i of solute in the slice is clearly:

$$w_i = c_i \, \Delta v_i \tag{16}$$

where Δv_i is the volume of the slice. Note that Δv_i is calculated from the elution time × the flow rate, and therefore requires that an accurate flow rate is known. The calculated mass W for the peak is then

$$W = \sum_{peak} w_i \tag{17}$$

where the sum is over the slices in the peak. When using this method, ASTRA calculates the peak mass according to Eq. (17) where the sum is over the slices in the peak within the limits set in the Peaks graph. Comparing this value with the injected mass for the peak, given an accurate injection volume and assuming that no part of the sample remains on the column(s), the calculated eluted mass should agree with the injected mass.

Known AUX Calibration Constant and 100% Mass Recovery

If you enable the "Assume 100% Mass Recovery" option in an experiment configuration, ASTRA assumes that 100% of the injected mass elutes in the peak area selected. In addition, the following information must be known:

- The total eluted mass for each peak (assumed to be equal to the injection mass).
- The detector's calibration constant, α.
- The flow rate with sufficient accuracy.

If you use this method, knowing the dn/dc is not required.

Using the known values, ASTRA calculates dn/dc by inverting Eq. (15) and using Eq. (18):

$$dn / dc = \frac{\alpha}{W_{injected}} \sum_{peak} \Delta v_i (V_i - V_{i,baseline}) \qquad (18)$$

You must be sure that all the injected mass elutes in the peak area selected and that the flow rate is accurate.

Determining Concentration by UV

This method is similar to determining concentration with a generic RI detector, except that the product of the extinction coefficient and the cell path length are used in place of dn/dc.

The value of alpha, called the UV response, may be calculated from the UV instrument's AUFS (absorption units at full scale) parameters and the maximum analog output value.

For example, some UV detectors have a maximum full scale output of 2 V. If the AUFS parameter on the UV instrument is set to 0.1 a.u. then alpha is (2 V / 0.1 a.u.) = 20 V/a.u. The calculated value should be entered into the Generic UV Instrument (UV) configuration field, as should the cell path length.

Determination of Molar Mass and Sizes

Different calculations can be used to determine molar mass and size depending on whether the solution is dilute or semi-dilute.

Dilute Solutions

At low concentrations typical of chromatographic separations, the distances between macromolecules are large enough that non-specific intermolecular interactions—otherwise known as "thermodynamic non-ideality"—may be ignored. This is generally true for concentrations below 0.1 - 1.0 mg/mL, though the cutoff concentration depends on the specific volume of the macromolecule and any long-range interactions due to poorly shielded charges, for example, with polyelectrolytes in low-ionic-strength buffers.

In the dilute, or ideal, limit, the relationship between the scattered light intensity, molar mass and concentration are given by Eq. (19):

$$\frac{R_\theta}{K^*} = Mc\,P(\theta) \qquad (19)$$

Where:

- R_θ is the excess Rayleigh ratio (cm^{-1}).
- K^* is an optical constant. $K^* = 4\pi^2 n_0^2 (dn/dc)^2 / (\lambda_0^4 N_A)$, where:
  - n_0 is the index of refraction of the solvent.
  - dn/dc is the differential refractive index increment of the solvent-solute solution with respect to a change in solute concentration, expressed in mL/g (this factor must be measured independently using a dRI detector).
  - λ_0 is the wavelength of the laser light in vacuum.
  - N_A is Avogadro's number, equal to 6.022 x 10^{23} mol^{-1}.
- M is the weight average molar mass (g/mol).
- c is the mass concentration of the solute molecules in the solvent (mg/mL).
- $P(\theta)$ is the theoretically-derived form factor, given by $P(\theta) = 1 - 2\mu^2 \langle r_g^2 \rangle / 3! + \cdots$, where $\mu = (4\pi/\lambda)\sin(\theta/2)$, and $\langle r_g^2 \rangle$ is the mean square radius. $P(\theta)$ is a function of the molecules' size, shape, and structure.

Eq. (19) is the basis for all molar mass and size calculations in ASTRA's chromatographic analyses when the Peaks view sets **A2** to zero and the **LS Analysis > Model** property to Debye. If the Zimm or Berry model is selected, then a variant of this equation is used (see page 390).

At each data slice, the angular signals are fit to this equation in order to determine M and r_g. $P(\theta)$ is expanded as a polynomial in $\sin^2(\theta/2)$ according to the value of **LS Analysis > Fit Degree** in the Peaks view.

Semi-Dilute Solutions

At higher sample concentrations, non-specific intermolecular interactions begin to have a first-order effect on the light scattering intensity. Again depending on properties such as the specific volume and degree of ionic shielding, the impact may be described by a first-order virial expansion in the form of Eq. (20)[1]:

$$\frac{K^* c}{R_\theta} = \frac{1}{MP(\theta)} + 2A_2 c \qquad (20)$$

Where A_2 is the second virial coefficient (mol mL / g^2).

In a Zimm plot (in batch mode), multiple concentrations are measured and Eq. (20) is fit to the complete data set $R_\theta(c)$ in order to obtain M, r_g, and A_2. ASTRA's calculation linearizes $1/P(\theta)$ as the following polynomial in $\sin^2(\theta/2)$ in order to avoid numerical instabilities.

$$\frac{1}{P(\theta)} \sim 1 + 2\mu^2 \left\langle r_g^2 \right\rangle / 3! - \ldots \qquad (21)$$

In chromatographic mode where each data slice is considered individually, the value of A_2 is assumed to be known (and entered into the sample parameters in the Peaks view), and Eq. (20) is fit to R_θ in order to obtain M and r_g.

ASTRA offers other models for analyzing molar mass and size, as well as A_2. Each of these models presents a different approximation to Eq. (20), which is formally identical to the result of thermodynamic fluctuation theory.

The Debye model, represented in Eq. (22), is optimal for determining r_g when A_2 may be ignored, while the Zimm model is optimal for determining A_2 when r_g can be ignored.

$$\frac{R_\theta}{K^* c} = MP(\theta) - 2A_2 c M^2 P^2(\theta) \qquad (22)$$

The Berry model, represented in Eq. (23), is implemented in ASTRA by linearizing $1/\sqrt{P(\theta)}$ to avoid numerical instability. It is a compromise between Zimm and Debye and is useful for many larger macromolecules:

$$\sqrt{\frac{K^* c}{R_\theta}} = \frac{1}{\sqrt{MP(\theta)}} + A_2 c \sqrt{MP(\theta)} \qquad (23)$$

In order to prevent ambiguous fitting results, ASTRA does not permit fitting to virial expansions beyond 2nd order.

1. B.H. Zimm, "The scattering of light and the radial distribution function of high polymer solutions," *J. Chem. Phys.*, vol. 16, pp. 1093-1099 (1948).

Conventional Method

First, construct a conventional plot. That is, plot $R_\theta / (K^* c)$ against $\sin^2(\theta/2)$. (This method is often referred to as the "Debye" method.) Next, fit a polynomial in $\sin^2(\theta/2)$ to the data for each of the m angles θ_i where measurements are collected for angles $i = 1,2,...,m$. From these fits, obtain the intercept, $R_0 /(K^* c)$, at $\theta = 0$ and the slope at zero angle, given by:

$$s = d\left[R_\theta / (K^* c)\right] \Big/ d\left[\sin^2(\theta/2)\right]_{\theta \to 0} \tag{24}$$

Note that as θ approaches zero, the form factor $P(\theta)$ approaches unity. Therefore, Eq. (19) becomes:

$$\frac{R_0}{K^* c} = M - 2A_2 c M^2 \tag{25}$$

If $A_2 = 0$, then

$$M = \frac{R_0}{K^* c} \tag{26}$$

On the other hand, solving Eq. (25) for M yields:

$$M = \frac{2\left(1 - \sqrt{1 - 8A_2 c\left(\frac{R_0}{K^* c}\right)}\right)}{8A_2 c} \tag{27}$$

Note that only one of the two solutions of Eq. (25) is physically reasonable. For very small values of A_2, we may rewrite Eq. (27) as follows:

$$M = \frac{2\left(\frac{R_0}{K^* c}\right)}{1 + \sqrt{1 - 8A_2 c\left(\frac{R_0}{K^* c}\right)}} \tag{28}$$

Eq. (28) is not susceptible to round-off error. Note that Eq. (28) reduces to $R_0 /(K^* c)$ as $A_2 \to 0$.

To find the mean square radius $\langle r_g^2 \rangle$ for the slice, we note that at very small angles $P(\theta) = 1 - 2\mu^2 \langle r_g^2 \rangle / 3! + \ldots$ and Eq. (19) may be written approximately as:

$$\begin{aligned} \frac{R_\theta}{K^* c} &= MP(\theta) - 2A_2 c M^2 P^2(\theta) \approx M\left[1 - 2\mu^2 \langle r_g^2 \rangle / 3!\right] - 2A_2 c M^2 \left[1 - 2\mu^2 \langle r_g^2 \rangle / 3!\right]^2 \\ &\approx M\left[1 - 2\mu^2 \langle r_g^2 \rangle / 3!\right] - 2A_2 c M^2 \left[1 - 4\mu^2 \langle r_g^2 \rangle / 3!\right] \end{aligned} \tag{29}$$

Eq. (24) may be written in terms of $\mu^2 = (4\pi/\lambda)^2 \sin^2(\theta/2)$. That is:

$$s = d\left[R_\theta/(K^*c)\right] \Big/ d\left[\sin^2(\theta/2)\right] = \frac{16\pi^2}{\lambda^2} d\left[R_\theta/(K^*c)\right] \Big/ d\left[\mu^2\right] \quad (30)$$

$$= -\frac{16\pi^2}{\lambda^2}\frac{M\left\langle r_g^2\right\rangle}{3}\{1-4A_2cM\}$$

Therefore:

$$\left\langle r_g^2\right\rangle = \frac{-3s\lambda^2}{16\pi^2 M(1-4A_2Mc)} \quad (31)$$

Zimm (Reciprocal) Method

To perform calculations with the Zimm method, which begins with a plot of $(K^*c)/R_\theta$ against $\sin^2(\theta/2)$, we expand the reciprocal of Eq. (19) to first order in c:

$$\frac{K^*c}{R_\theta} = \frac{1}{MP(\theta)} + 2A_2c \quad (32)$$

By following the procedures described for the conventional plot, we obtain the following results:

$$M = \left(\frac{K^*c}{R_0} - 2A_2c\right)^{-1} \quad (33)$$

and

$$\left\langle r_g^2\right\rangle = \frac{3sM\lambda^2}{16\pi^2} \quad (34)$$

where:

$$s = d\left[K^*c/R_\theta\right] \Big/ d\left[\sin^2(\theta/2)\right]_{\theta\to 0} \quad (35)$$

Berry (Square Root) Method

To perform calculations with the Berry method, which begins with a plot of $\sqrt{K^* c / R_\theta}$ against $\sin^2(\theta/2)$, we must expand the square root of the reciprocal of Eq. (19) to first order in c:

$$\sqrt{\frac{K^* c}{R_\theta}} = \frac{1}{\sqrt{MP(\theta)}} + A_2 c\sqrt{MP(\theta)} \tag{36}$$

In this case the results are:

$$M = \frac{4}{\left(\sqrt{K^* c / R_\theta} + \sqrt{K^* c / R_\theta - 4A_2 c}\right)^2} \tag{37}$$

and

$$\left\langle r_g^2 \right\rangle = \frac{3\lambda^2 S}{8\pi^2 \sqrt{M}(1/M - A_2 c)} \tag{38}$$

where:

$$S \equiv \frac{d\left[\sqrt{K^* c / R_\theta}\right]}{d\left[\sin^2(\theta/2)\right]_{\theta \to 0}} \tag{39}$$

Assuming a Molecular Structure

For a variety of possible molecular structures, the interpretation of measurements may be simplified considerably if such structures are known in advance. Of course, if the scattering molecules are not of the assumed structure, significant errors can result. The key to their use, of course, requires the applicability of the Rayleigh Gans approximation (see “Number Density Calculation” on page 412). In that event, knowing the molecular structure in advance often results in an exact analytical expression for the form factor $P(\theta)$.

Consider first an assumed Random Coil structure. Returning again to Eq. (19), we replace the theoretical form factor $P(\theta)$ by its exact form first derived by Debye[1]:

$$P(\theta) = \frac{2}{u^2}\left(e^{-u} - 1 + u\right) \tag{40}$$

where $u = (4\pi/\lambda^2)\left\langle r_g^2 \right\rangle \sin^2(\theta/2)$.

1. P. Debye, “Molecular-weight determination by light scattering,” *J. Phys. Coll. Chem.*, vol. 51, pp. 18-32 (1947).

Since $P(\theta)$ is a nonlinear function of the mean square radius $\langle r_g^2 \rangle$, we must use an iterative nonlinear least squares fit of the data to this model. Unlike the other methods, the Random Coil method assumes the polymers are approximately random coils. This can be an advantage for large random coil molecules because it allows the fit to proceed with fewer parameters than would otherwise be required in a simple polynomial fit, and the result can be lower estimated errors.

For the Peaks procedure (page 221), you can select which calculation type—Zimm, Berry, Debye, or random coil—you wish to employ. If you own a miniDAWN, the calculation type must be either Zimm or Random Coil.

The result of these calculations is that for each slice i we have the molar mass M_i and the mean square radius $\langle r_g^2 \rangle_i$.

Assuming good chromatographic separation, these quantities can be used together with the concentration c_i (measured with a concentration-sensitive detector) to find the molar mass and radius moments, as described next.

Other Structures Method

For certain other forms of molecular structures (for example, sphere, coated sphere, and rod), we fit the Zimm equation to $R_\theta / K^* c$ vs. $\sin(\theta/2)$. As in the Conventional method, we insert into Eq. (19) the theoretical form factor $P(\theta)$ for the desired model. Form factor models have been derived for spheres, coated spheres, and rods and are covered in the text by van de Hulst[1]. Note that the sphere and coated sphere models yield radii, while the rod model produces a length.

Spheres

$$P(\theta) = \frac{3}{u^3}\left(\sin u - u\cos u\right) \qquad (41)$$

where $u = (4\pi r / \lambda)\ \sin(\theta/2)$.

Rods

$$P(\theta) = \left(\frac{1}{u}\right)\int_0^{2u} \frac{\sin t}{t}\,dt - \frac{\sin^2 u}{u^2} \qquad (42)$$

where $u = [(2\pi n_o / \lambda_o)\, L\ \sin(\theta/2)]$, and L is the rod length, which is assumed to be much greater than its negligible diameter.

1. H.C. van de Hulst, *Light Scattering by Small Particles*, Wiley, New York (1957)

Molar Mass and rms Radius Moments

In chromatographic mode, ASTRA calculates the following molar mass and rms (root mean square) radius moments for each peak selected. Naturally, moments may be referenced to averages over the entire sample, which may include many peaks.

Number-average molar mass:

$$M_n = \frac{\sum_i n_i M_i}{\sum_i n_i} = \frac{\sum_i c_i}{\sum_i c_i / M_i} \tag{43}$$

Note that an ASTRA measurement usually requires an independent concentration determination. Since the relation between concentration (mg/mL) and number density (number/mL) is simply $nM = c$, the results of Eq. (43) follow immediately.

Weight-average molar mass:

$$M_w = \frac{\sum_i n_i M_i^2}{\sum_i n_i M_i} = \frac{\sum_i c_i M_i}{\sum_i c_i} \tag{44}$$

z-average molar mass:

$$M_z = \frac{\sum_i n_i M_i^3}{\sum_i n_i M_i^2} = \frac{\sum_i c_i M_i^2}{\sum_i c_i M_i} \tag{45}$$

The measurement of the mean square radius, $\langle r_g^2 \rangle$, by light scattering invariably requires measurement of the product of the molar mass times this quantity. The result depends also upon the concentration of the molecules. Thus measurement of the effective mean square radius is weighted by cM. Accordingly, we derive for each peak selected a scattered light weighting as:

$$\langle r_g^2 \rangle_{LS} = \frac{\sum_i c_i M_i \langle r_g^2 \rangle_i}{\sum_i c_i M_i} \tag{46}$$

This quantity is usually referred to as the z-average mean square radius, $\langle r_g^2 \rangle_z$, though this definition is quite strange. Specifically, it arises from the polymer chemistry nomenclature for a so-called ideal random coil structure whereby the molar mass is directly proportional to the mean square radius to the 0.5 power, that is:

$$M = a\langle r_g^2 \rangle^{0.5} \text{ or } \langle r_g^2 \rangle \propto M^2 \tag{47}$$

where a is a constant. Substituting this value of $\langle r_g^2 \rangle$ into Eq. (46) yields:

$$\langle r_g^2 \rangle_{LS} = \frac{\sum_i c_i M_i M_i^2}{a^2 \sum_i c_i M_i} = \frac{\sum_i c_i M_i^3}{a^2 \sum_i c_i M_i} \tag{48}$$

But this is identical (except for the constant a^2) to Eq. (45), the so-called z-average molar mass. This is the origin of the light scattering derived value of the mean square radius, that is:

$$\langle r_g^2 \rangle_z \equiv \langle r_g^2 \rangle_{LS} \tag{49}$$

Defining Eq. (46) as the z-average mean square radius suggests that there are number-average and weight-average possibilities, as well. These are reported by ASTRA, although their actual significance is not clear.

Number-average mean square radius:

$$\langle r_g^2 \rangle_n = \frac{\sum_i \frac{c_i}{M_i} \langle r_g^2 \rangle_i}{\sum_i \frac{c_i}{M_i}} \tag{50}$$

Weight-average mean square radius:

$$\langle r_g^2 \rangle_w = \frac{\sum_i c_i \langle r_g^2 \rangle_i}{\sum_i c_i} \tag{51}$$

The quantities c_i, M_i, and $\langle r_g^2 \rangle_i$ in these equations are respectively the mass concentration, molar mass (g/mol), and mean square radius of the i^{th} slice. The often referenced root-mean-square radii (rms) are simply the square roots of the associated mean square radii.

ASTRA also calculates two polydispersity values: $\rho = M_w / M_n$ and $\rho' = M_z / M_w$. Only the former is found in the literature.

All measurements processed by ASTRA are weighted by the standard deviations of the measured quantity. These standard deviations are then used to generate the expected standard deviations of all derived quantities. The uncertainty of the weight-average molar mass (M_{avg}) is then calculated as follows:

$$M_{w-avg} = \frac{\sum_i M_i \frac{1}{\sigma_{M_i}^2}}{\sum_i \frac{1}{\sigma_{M_i}^2}} \qquad (52)$$

σ_{M_i} is the uncertainty in the value of M_i. The error in this calculation is defined as follows:

$$\sigma_{M_{w-avg}} = \frac{1}{\sum_i \sigma_{M_i}^2} \qquad (53)$$

Uncertainties in Calculated Quantities

ASTRA calculates uncertainties for all reported quantities. By analyzing the baseline data at the beginning and end of the chromatogram, ASTRA determines the statistical fluctuation in each detector's output, including all photodiodes and the AUX signals.

Each detector is weighted based on the fluctuations (noise) seen in the first and last 10% of the data points, up to 100 data points. Whichever end is least noisy is used to calculate the weighting factor. (For batch mode calculations, data points within each plateau are used to calculate the detector weighting factors for each concentration.)

The error bars in the analysis plot do not represent this weighting factor directly. The analysis plot involves performing an nth order polynomial fit to $R_\theta / K^* c$ (for the Conventional Method), $K^* c / R_\theta$ (for the Zimm (Reciprocal) Method), $\sqrt{K^* c / R(\theta)}$ (for the Berry (Square Root) Method), or $P(\theta)$ (for the Other Structures Method). The error bar calculation therefore involves the weighting factor, the normalized R_θ value as well as a concentration uncertainty factor and the Chi-squared value returned from the fit. If the normalization is off for some detectors, then the Chi-squared value from the fit tends to increase, causing all error bars to grow. Hence, changes to the normalization coefficients will affect the error bars shown in the analysis plot, as well as the uncertainties in the overall peak results.

The different errors combine according to the usual rules for propagation of errors to yield a standard deviation (depending on calculation method) for each slice. These in turn allow calculation of uncertainties in the molar mass and size for each slice, and hence uncertainties in the calculated molar mass and size averages.

Remember, these uncertainties are statistical only, and do not include any of the many possible systematic errors that may be present. Examples are errors in dn/dc, the DAWN calibration constant, the AUX calibration constants, and the normalization coefficients.

Use the reported uncertainties as a measure of the statistical consistency of the data, never as an absolute limit on the error in your results.

Out of Range Values

Occasionally, electrical noise or a very low concentration or light scattering signal may cause the calculated molecular weight at a particular slice to be a negative number. For low molecular weights, often the mean square radius at a particular slice will be negative due to random noise in the analysis plot for that slice. Also, noise may cause both the calculated molecular weight and the mean square radius to have uncertainties larger than the values themselves. In these cases, special considerations are called for.

When calculating molecular weight averages, ASTRA first checks the calculated molecular weight values of all slices to be included in the calculations to find out if any of them are negative. ASTRA then removes slices that have negative values before calculating the averages.

When calculating mean square radius averages ASTRA includes values from all slices in the summation. If the sum of the mean square radii is positive, ASTRA will calculate the root mean square averages. If it is negative, the resulting root mean square averages will be set to zero.

In addition to the above, if any of the slices to be included in the averages have uncertainties larger than the values themselves, ASTRA will exclude them from the averages.

When plotting data in the Distribution Plots, ASTRA removes any slices that have negative values.

Differential Distribution Calculations

ASTRA 7 uses an adaptive binning technique for determining the differential distributions. It works both with the direct results, and with data that has been fit with results fitting.

Branching Calculations

ASTRA performs a number of sophisticated branching calculations. These are described below.

Branching Ratio: Radius Method

The branching ratio g_M is formally defined[1] as:

$$g_M = \left(\frac{\langle r^2 \rangle_{br}}{\langle r^2 \rangle_{lin}} \right)_M \qquad (54)$$

where $<r^2>_{br}$ and $<r^2>_{lin}$ are the mean square radii of branched and linear (unbranched) polymer samples to be compared. Note that the ratio is taken at the same molar mass, *not* at the same volume. In general, for a given molar mass, the branched polymer will have a smaller radius, so g_M will lie between 0 and 1.

ASTRA calculates g_M this way: If no results fitting method has been selected, ASTRA uses the raw rms radius vs. molar mass data for both the linear and branched files. If a results fitting method has been selected, ASTRA uses the fitted data from molar mass vs. volume and rms radius vs. volume directly in the branching calculations. For a number of points (300 points per decade of molar mass), Eq. (54) is applied.

In order to obtain useful branching information, the two files (linear and branched) should overlap as much as possible in molar mass. The branching ratio g_M can only be calculated in this region of overlap, since only in this region can radii be found at the same molar mass.

To use this method, select the Radius method in the Branching properties view. See “Branching Procedure” on page 253.

Branching Ratio: Mass Method

If the molecular radii are too small to be calculated accurately, then we must use another method. Assuming the Flory-Fox equation is valid,[2] it can be shown that:

$$g_M = \left(\frac{M_{lin}}{M_{br}} \right)_V^{\frac{(a+1)}{e}} \qquad (55)$$

1. B.H. Zimm and W.H. Stockmayer, “The dimensions of chain molecules containing branches and rings,” *J. Chem. Phys.*, vol. 17, pp. 1301-1314 (1949).
2. L.P. Yu and J.E. Rollings, “Low-angle light scattering-aqueous size exclusion chromatography of polysaccharides: Molecular weight distribution and polymer branching determination,” *J. Appl. Polym. Sci.*, vol. 33, pp. 1909–1921 (1987).

where M_{lin} and M_{br} are the molar masses of a linear and branched polymer, respectively, α is the Mark-Houwink-Sakurada parameter for the linear polymer, and e is the drainage parameter, ranging from 0.5 for a non-draining polymer to 1.0 for a free-draining polymer to 1.5 for a Flory-Fox polymer.[1] A value of 0.5–1.0 seems most used in the literature. The effect of the choice of e on the results can be seen in the figure below, which shows g_M for various values of e using a Mark-Houwink-Sakurada parameter α of 0.7, typical for a random coil. Notice that the ratio in Eq. (55) is taken at constant elution volume V.

If no results fitting method has been selected, ASTRA uses the raw molar mass vs. volume data for both the linear and branched files. If a results fitting method has been selected, ASTRA uses the fitted data from molar mass vs. volume in the branching calculations. For each slice of the branched file, ASTRA obtains the linear molar mass from the slice in the linear file having the elution volume closest to that of the branched slice.

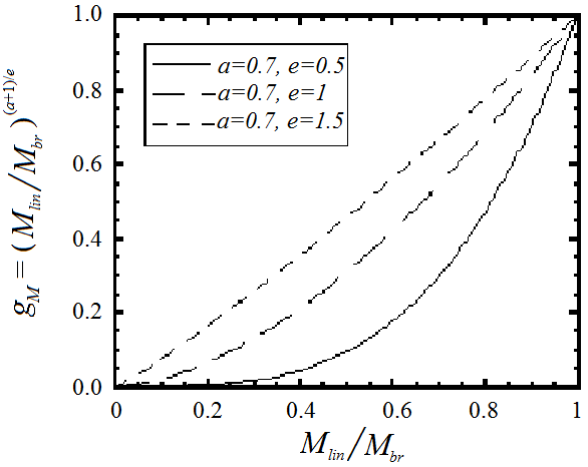


Figure E-1: Branching ratio g_M as a function of α and e for the Mass Method

For this method, the volumes should have a large region of overlap for an effective plot. To use this method, select the Mass method in the Branching properties view for each branched file. See "Branching Procedure" on page 253.

Branching Ratio: Viscosity Method

The branching ratio $g'M$ is formally defined as:

$$g'(M) = \frac{[\eta(M)]_{br}}{[\eta(M)]_{lin}} \qquad (56)$$

1. B.H. Zimm and R.W. Kilb, "Dynamics of branched polymer molecules in dilute solution," *J. Polym. Sci.,* vol. 37, pp. 19-42 (1959).

where η_{br} and η_{lin} are the intrinsic viscosities of branched and linear (unbranched) polymer samples to be compared. Note that the ratio is taken at the same molar mass, *not* at the same volume. In general, for a given molar mass, the branched polymer will have a smaller radius, so $g'M$ will lie between 0 and 1.

The Radius and Viscosity branching ratios are related by a term called the "drainage parameter" (e), as follows:

$g^e = g'$

ASTRA uses the raw intrinsic viscosity vs. molar mass data for both the linear and branched files. For a number of points (300 points per decade of molar mass), Eq. (56) is applied.

In order to obtain useful branching information, the two files (linear and branched) should overlap as much as possible in molar mass. The branching ratio $g'M$ can only be calculated in this region of overlap, since only in this region can intrinsic viscosity be found at the same molar mass.

To use this method, select the Viscosity method in the Branching properties view. See "Branching Procedure" on page 253.

Branching Per Molecule

The number of branch points per molecule is related to the branching ratio, but some knowledge of the type of branching is necessary. You can choose either trifunctional (Y or T) or tetrafunctional (X) branching, and monodisperse or polydisperse slices. You can also choose comb and star branching models, but they are only available for monodisperse distributions.

These formulas[1] relate g_M to B *for randomly branched polymers:*

Trifunctional Branching

- Polydisperse:

$$g_M = \frac{6}{B_{3w}}\left\{\frac{1}{2}\left(\frac{2+B_{3w}}{B_{3w}}\right)^{1/2}\ln\left[\frac{\left(2+B_{3w}\right)^{1/2}+B_{3w}^{1/2}}{\left(2+B_{3w}\right)^{1/2}-B_{3w}^{1/2}}\right]-1\right\} \qquad (57)$$

- Monodisperse:

$$g_M = \left[\left(1+\frac{B_{3n}}{7}\right)^{1/2}+\frac{4B_{3n}}{9\pi}\right]^{-1/2} \qquad (58)$$

1. B.H. Zimm and W.H. Stockmayer, *ibid.*

Tetrafunctional Branching

- Polydisperse:

$$g_M = \frac{\ln\left(1 + B_{4w}\right)}{B_{4w}} \tag{59}$$

- Monodisperse:

$$g_M = \left[\left(1 + \frac{B_{4n}}{6}\right)^{1/2} + \frac{4B_{4n}}{3\pi}\right]^{-1/2} \tag{60}$$

Star Branching

- Monodisperse[1]:

$$g_M = \frac{6B}{B^2 + 3B + 2} \tag{61}$$

Comb Branching

- Monodisperse[2]:

$$g_M = \frac{2B^3 + 12B^2 + 10B + 3}{4B^3 + 12B^2 + 11B + 3} \tag{62}$$

For each of these relations, the left hand side, g_M, is known already (see the previous sections). The appropriate equation is solved for B for each slice which produced a reasonable value of g_M. Note that if g_M falls outside the range 0 to 1, no value of B will be calculated for that slice.

The legend is labeled with the specified functionality: "3" for Trifunctional branching or "4" for Tetrafunctional branching; "n" for Monodisperse slices or "w" for Polydisperse. For each branched file to be plotted, select the branching functionality and whether the slices are monodisperse or polydisperse in the Branching property view.

The next figure, a plot of Eqs. (41)–(44), shows how the B's are related to g_M for the various branching options. Note that these relations assume randomly branched polymers. Also note that different assumptions about

1. B.H. Zimm and W.H. Stockmayer, *ibid*.
2. W. Radke, A. H. E. Müller, "Synthesis and Characterization of Comb-Shaped Polymers by SEC with On-Line Light Scattering and Viscometry Detection", *Macromolecules*, vol. 38, pp. 3957. (2005)

functionality and dispersity yield quite different values of B for the same value of g_M. Thus some knowledge of the type of branching is necessary for a plot of branches per molecule to have any meaning.

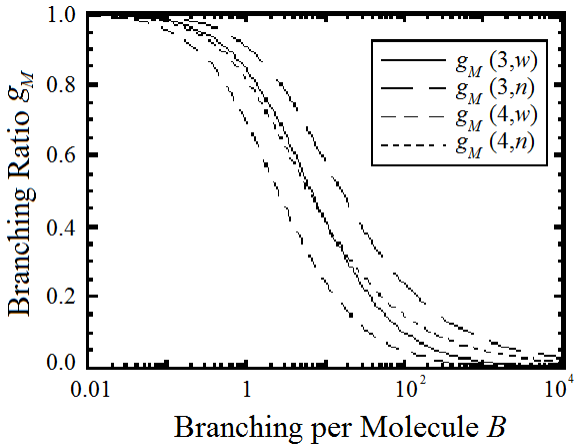


Figure E-2: Branching ratio g_M as a function of B for various branching options

Long Chain Branching

The long chain branching per 1000 repeat units is defined for each slice as

$$\lambda = 1000B\frac{R}{M} \qquad (63)$$

where B is the branching per molecule for the slice (as calculated above), R is the repeat unit molar mass, and M is the branched molar mass for the slice. You must enter the repeat unit molar mass in the Unit MW box in the Branching property view for each branched file to be plotted.

F Particles

Particles support is an add-on option for ASTRA. This option provides a procedure to calculate particle number densities (see "Number from LS Procedure" on page 242). Particles include colloidal particles such as polystyrene latex spheres, liposomes, and vesicles. Particle measurements are especially suited for use when a light scattering instrument is coupled to a fractionation technique such as Field Flow Fractionation (FFF).

CONTENTS **PAGE**

Determination of Sizes 406
Radius Moments Calculation 409
Number Density Calculation 412

Determination of Sizes

As discussed in Appendix E, "Light Scattering Theory", data collected by a DAWN system are used to derive molecular parameters. In the ASTRA Particles mode, however, concentration is not measured and dn/dc need not be known in order to determine size. It is assumed further that the particle concentrations are low enough that there are no interactions of consequence between them, and the 2nd virial coefficient is assumed to be zero. Most importantly, it is assumed that the particles present have been fractionated, so that each slice contains particles of identical size.

Only measurement of the variation of the scattered intensity with angle is made, that is, only the excess Rayleigh ratio need be measured. This angular variation is then matched to an assumed particle shape to obtain a corresponding size. If nothing is known of the particle shape, for certain classes of particles it may still be possible to obtain a measure of the particle's mean square radius. Interpretation of these values in terms of explicit particle size characteristics again requires some assumed model.

The particles considered, except those analyzed by means of the Lorenz-Mie theory (discussed in "Homogeneous Spheres Using Exact Lorenz-Mie Theory" on page 408) must satisfy the Rayleigh-Gans criteria:

$$\left|m-1\right| << 1, \text{ where } m = n / n_0 \text{ and}$$
$$ka\left|m-1\right| << 1, \text{ where } k = 2\pi / \lambda \qquad (1)$$

and a is a characteristic dimension of the particle.

Conventional Mean Square Radius Determination

First, construct a conventional plot excluding the optical constant and concentration terms, that is, create a plot of R_θ vs. $\sin^2(\theta/2)$. Second, fit a polynomial in $\sin^2(\theta/2)$ to the data, and thereby obtain the intercept at zero angle, R_0, as well as the slope at zero angle, $s = d[R_\theta] / d[\sin^2(\theta/2)]_{\theta=0}$. Thus we have the following, which is in agreement with Appendix E.

$$\left\langle r_g^2 \right\rangle = \frac{-3s\lambda^2}{16\pi^2 R_0} \qquad (2)$$

In the Zimm, Debye, and Berry models, the mean square radius r_g is determined by fitting Eq. (20), Eq. (22) or Eq. (23) respectively to the angular data.

Note that this result is valid only in the Rayleigh-Gans limit wherein the excess Rayleigh ratio is assumed to be proportional to $P(\theta)=1-2\mu^2\left\langle r_g^2\right\rangle/3!+\cdots$ as given in "Determination of Molar Mass and Sizes" on page 389.

Mean Square Radius Determination from an Assumed Random Coil

As discussed in "Determination of Molar Mass and Sizes" on page 389, we insert the theoretical form factor $P(\theta)$ for random coils into Eq. (19).

$$P(\theta) = \frac{2}{u^2}\left(e^{-u} - 1 + u\right) \tag{3}$$

where $u = (4\pi / \lambda^2) \left\langle r_g^2 \right\rangle \sin^2(\theta / 2)$

Since $P(\theta)$ is a nonlinear function of its parameter, $\left\langle r_g^2 \right\rangle$, we use an iterative nonlinear least squares fit to the Zimm formalism. Unlike the other fit methods, the Random Coil method assumes the polymers are random coils. This can be an advantage for large random coil molecules, because it allows the fit to proceed with fewer parameters than would otherwise be required in a simple polynomial fit, and the result can be lower estimated errors. The only size derived, of course, is the mean square radius. For a polymer comprised of N segments of length a, the relationship between r_g, N, and a of such a polymer in a theta solvent is given by:

$$\left\langle r_g^2 \right\rangle = Na^2 / 6 \tag{4}$$

Size from a Known Structure

To derive a particle size based on a structure known *a priori*, we again plot R_θ vs. $\sin^2(\theta / 2)$ and replace the theoretical form factor $P(\theta)$ by the appropriate model assumed. Appropriate form factors have been derived for spheres, coated spheres, and rods. They are covered in the text by van de Hulst[1]. Note that the sphere and coated sphere models yield a radius, while the rod model produces a length.

Sphere

$$P(\theta) = \left[\frac{3}{u^3}\left(\sin u - u\cos u\right)\right]^2 \tag{5}$$

where $u = 2ka \sin(\theta/2)$.

Rod

$$P(\theta) = \left(\frac{1}{u}\right)\int_0^{2u} \frac{\sin t}{t}\,dt - \frac{\sin^2 u}{u^2} \tag{6}$$

where $u = [(2\pi n_o / \lambda_o) L \sin(\theta/2)]$, and L is the rod length, where L is assumed to be much greater than the rod diameter.

1. H.C. van de Hulst, *Light Scattering by Small Particles*, Wiley, New York (1957)

Homogeneous Spheres Using Exact Lorenz-Mie Theory

If particles are known to be homogeneous spheres, their radii may be derived from the exact scattering theory developed by Ludvig Lorenz and referred to as the Lorenz-Mie theory. This theory represents an exact solution of Maxwell's electromagnetic theory. There are no restrictions on the particle's refractive index or size, so the ASTRA software may be used to determine the radius of homogeneous spherical particles including latex spheres and even gold and carbon particles whose refractive indices are complex.

From the Lorenz-Mie theory, the measured values of R_θ at the angular set measured are used to extract the radius producing the best fit to the theory in a least squares sense.

The extended Lorenz-Mie exact calculation for a coated sphere (single layer) is given in the text by Bohren and Huffman based on the paper of A. L. Aden and M. Kerker[1].

1. A. L. Aden and M. Kerker, "Scattering of electromagnetic waves from two concentric spheres," *J. Appl. Phys.,* vol. 22, pp. 1242-1246 (1951)

Radius Moments Calculation

ASTRA calculates the mean square radius moments for each peak as discussed in "Molar Mass and rms Radius Moments" on page 395. Although only the LS average (also referred to as the z-average) is most commonly measured, the other two are shown for completeness. The specific type of radius (radius, rms radius, or hydrodynamic radius) depends on the type of analysis being performed. As in Appendix E, all summations are taken over one peak.

rms Radius

LS Average

$$\left\langle r_g^2 \right\rangle_{LS} = \frac{\sum_i R_{0_i} \left\langle r_g^2 \right\rangle_i}{\sum_i R_{0_i}} \equiv \left\langle r_g^2 \right\rangle_z \tag{7}$$

The quantities R_0, and $\left\langle r_g^2 \right\rangle_i$ in Eq. (7) and V_i in the equations that follow are respectively the Rayleigh Ratio, mean square radius, and volume of the i^{th} slice. The root mean square (rms) radii are simply the square roots of the appropriate mean square radii.

Number Average

$$\left\langle r_g^2 \right\rangle_n = \frac{\sum_i \frac{R_0 \left\langle r_g^2 \right\rangle_i}{V_i^2}}{\sum_i \frac{R_0}{V_i^2}} \tag{8}$$

Weight Average

$$\left\langle r_g^2 \right\rangle_w = \frac{\sum_i \frac{R_0 \left\langle r_g^2 \right\rangle_i}{V_i}}{\sum_i \frac{R_0}{V_i}} \tag{9}$$

Uncertainty Weighted Average

$$\left\langle r_g^2 \right\rangle_{avg} = \frac{\sum_i \left\langle r_g^2 \right\rangle_i \frac{1}{\sigma^2_{\left\langle r_g^2 \right\rangle_i}}}{\sum_i \frac{1}{\sigma^2_{\left\langle r_g^2 \right\rangle_i}}} \quad (10)$$

Where $\langle r^2 \rangle_i$ is as defined previously, and $\sigma^2_{\left\langle r_g^2 \right\rangle_i}$ is the uncertainty in the mean square radius measurement. The error in this calculation is defined as follows:

$$\sigma^2_{\left\langle r_g^2 \right\rangle_{avg}} = \frac{1}{\sum_i \sigma^2_{\left\langle r_g^2 \right\rangle_i}} \quad (11)$$

Radius and Hydrodynamic Radius

The quantities R_0, r_i, and V_i in these equations are respectively the Rayleigh Ratio, radius (either radius or hydrodynamic radius), and the volume of the i^{th} slice.

Number Average

$$R_n = \frac{\sum \frac{R_0 r_i}{V_i^2}}{\sum \frac{R_0}{V_i^2}} \quad (12)$$

Weight Average

$$R_w = \frac{\sum \frac{R_0 r_i}{V_i}}{\sum \frac{R_0}{V_i}} \quad (13)$$

Z-Average

$$R_z = \frac{\sum R_0 r_i}{\sum R_0} \quad (14)$$

Uncertainty Weighted Average

$$r_{avg} = \frac{\sum r_i \frac{1}{\sigma_{r_i}^2}}{\sum \frac{1}{\sigma_{r_i}^2}} \quad (15)$$

where r_i is as defined previously, and σ_{r_i} is the uncertainty in the radius measurement. The error in this calculation is defined as follows:

$$\sigma_{r_{avg}} = \frac{1}{\sum \sigma_{r_i}^2} \quad (16)$$

Number Density Calculation

This section discusses how the distribution plots are calculated and why a model is needed.

The mean square radius is given by Eq. (17), where the distances r_i are measured from the particle's center of mass to the mass element m_i.

$$< r^2 >= \frac{\sum_i r_i^2 m_i}{\sum_i m_i} = \frac{1}{M}\int r^2 dm \qquad (17)$$

Eq. (17) refers to a single particle whereas the quantity actually measured from an ensemble of particles may be shown to be a so-called LS-average mean square radius.[1] Were the particles random coils in a theta solvent, then this would be the so-called z-average mean square radius. We assume that the particle size distribution within each slice of an eluting sample following separation is essentially monodisperse. Therefore the particles in slice i, each of mass M_i, are assumed to have the same mean square radius. We define the root mean square radius as the square root of the mean square radius or, simply, $r_g = < r^2 >^{1/2}$.

The Rayleigh-Gans-Debye approximation (RGD):

$$\frac{K^* c}{R(\theta)} \approx \frac{1}{M_w P(\theta)} \qquad (18)$$

can be re-written in the limit as $\theta \to 0$, P(0) = 1, and we have:

$$R(0) = K^* c_i M_i = K^* n_i M_i^2 \qquad (19)$$

since the concentration of mass in the ith slice is $c_i = n_i M_i$. If the elements of the particle whose molar mass is M_i are of uniform density and occupy a volume V_i, then the number of particles per mL in the ith slice, n_i, is proportional to the extrapolated zero-angle Rayleigh ratio divided by the square of the particle's volume, i.e.:

$$n_i \propto R(0) / V_i^2 \qquad (20)$$

Therefore we can write the *number fraction* of particles within slice i as n_i / D where:

$$D = \sum_j n_j \qquad (21)$$

1. P.J. Wyatt, "New Insights into GPC Combined with MALS," *Waters Corporation GPC Symposium Proceedings* (San Diego, 1996).

is the summation taken over all slices in the selected region (or peak) of the eluting fractions. Note that although M_i is the *molar* mass of the particles, that value is proportional to the *mass* of the particles. Both are proportional to the *volume* of the particles if the volume is of uniform density.

Although the analysis of each slice results in a corresponding value of r_g, there may be other slices with similar sizes due to experimental fluctuations in the derived values. The expected monotonic variation of r_g with elution volume may be obtained by fitting the calculated values to a selected functional form using a least squares procedure. Alternatively, the slice data may be sorted into a set of size bins to obtain the differential number fraction after dividing each such fraction by the bin size. The fractions may also be distributed over the range of size bins included within the measured standard deviation associated with the particular contributing fraction.

In any event, the differential number fraction $n(r)\ dr$ of particles in the selected peak region between r and $r + dr$ now may be calculated explicitly without any advance knowledge of the mass concentration at each slice provided we know the particle structure and that the RGD approximation is valid. For example, if we know that the particles are homogeneous spheres, we may replace V_i^2 by r_g^6. There are many other particle shapes where the relation between r_g and V_i^2 is known. The differential mass fractions may be generated in a similar manner without reference to a second detector.

What about particles whose shape is not known *a priori*? Although we may still calculate r_g as a function of elution volume (the r_g "calibration curve"), we cannot determine the differential number or mass fractions. Indeed, if we do not know the relation between the measured r_g and the particle's hydrodynamic radius, we cannot generate differential distributions. Were we to add a concentration detector following the LS detector, we could easily generate the differential mass fraction distributions of r_g.

A few other points must be discussed; most important among them is the applicability of the RGD approximation assumed in the preceding analysis. The simplest particles most frequently measured by particle sizing procedures are the polystyrene latex (PSL) spheres (emulsions) whose refractive index at wavelengths in the visible is about 1.59. Relative to water, whose refractive index is about 1.33, these spheres have a relative refractive index m = 1.59/1.33 $\approx$ 1.2. Rigorous application of the RGD theory requires that m - 1 << 1, which is a slight stretch for these PSL spheres. Perhaps more importantly, the phase shift of a wave passing through the particle, $2\pi a[m\text{-}1]\ n_0/\lambda_0$, where a is the sphere radius, also must be << 1. Even if we make the assumption that 0.2 << 1, attempting to size larger submicron particles using this approximation will quickly lead us out of the range of RGD applicability!

The saving grace of this approach is twofold: first, the theory happens to work significantly better than one might expect, even when the RGD requirements are not strictly satisfied, and second, the pertinent values are calculated in the limit $\theta \rightarrow 0$ (as shown in Eq. (19)), a regime where the RGD requirements are much more easily satisfied. As the scattering angle becomes very small, the RGD approximation becomes more valid as was confirmed vividly by the analyses of Kerker et al.[1] The result is that values of R_0 may be generated directly from the measurements if the particle's structure is known, or from the more general expansion of the form factor $P(\theta)$.[2]

Furthermore, many of the calculated results for the analyses of distributions of PSL spheres may be checked with more exact LS theory to confirm the precision of the sizes measured using the RGD approximation. Applying the Lorenz-Mie theory confirms the results derived by the present treatment. In addition, average values measured by photon correlation spectroscopy (PCS) at individual slices also confirm the average values generated by the present implementation of RGD theory.

1. M. Kerker, W.A. Farone, and E. Matijevic, "Applicability of Rayleigh-Gans Scattering to Spherical Particles," *J. Opt. Soc. Am.,* vol. 53, pp. 758-759 (1963).
2. P.J. Wyatt, "Light scattering and the absolute characterization of macromolecules," *Analytica Chimica Acta*, vol. 272, pp. 1-40 (1993).

G QELS Theory

This appendix gives a quick overview of the theory behind cumulants and regularization, which are analysis techniques used with QELS data. This includes descriptions of the implementation in ASTRA and interpretation of results.

CONTENTS **PAGE**

Cumulants Theory 416
Regularization Theory 421

Cumulants Theory

The analysis of QELS data is straightforward for a monodisperse sample. For unfractionated, polydisperse samples, however, the analysis becomes much more complicated. The simplest approach to analyzing data from polydisperse samples is to assume that the sample is monodisperse, apply the analysis from ASTRA, and come up with some sort of mass-averaged result for the hydrodynamic radius. The measured correlation function for a polydisperse sample actually contains additional information, and several strategies have been developed to extract more information about the underlying size distribution from the correlation function.

The next level of sophistication in QELS analysis for polydisperse, unfractionated samples is the method of cumulants. In a nutshell, the method of cumulants involves fitting the correlation function not to a single decay time, but to a Gaussian distribution of decay times. The method of cumulants retrieves the mean and variance for this distribution.

Cumulants Method Computations

The result of a QELS measurement is a second order correlation function:

$$g^{(2)}(\tau) = \frac{\langle I(t)I(t+\tau)\rangle}{\langle I(t)\rangle^2} \qquad (1)$$

where $I(t)$ is the intensity of the scattered light at time t, and the brackets indicate averaging over all t. The correlation function depends on the delay τ, that is, the amount that a duplicate intensity trace is shifted from the original before the averaging is performed. A typical correlation function for a monodisperse sample is shown in Figure G-1.

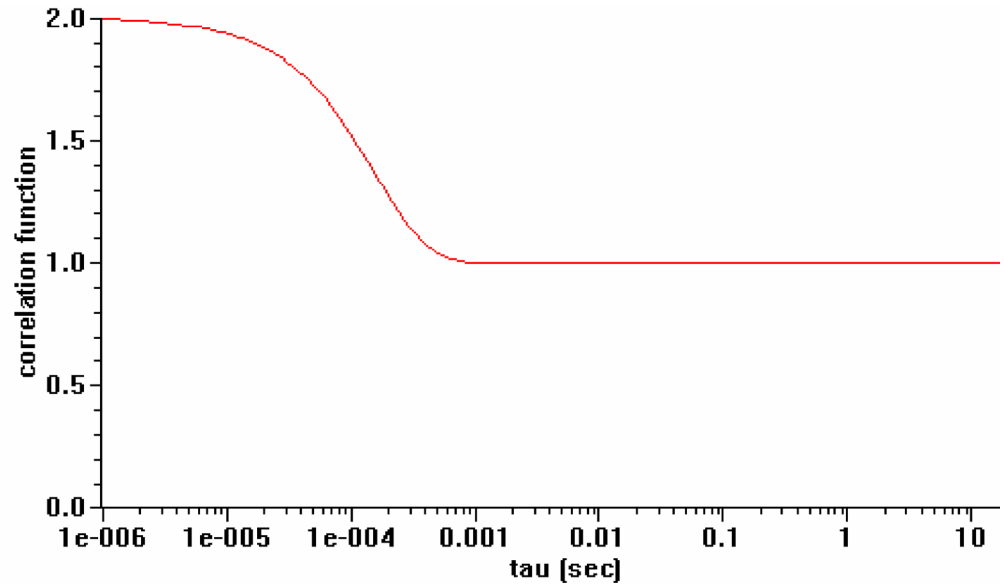


Figure G-1: Correlation function for a multi-tau correlator like that in WyattQELS

As described in various light scattering texts[1], the correlation function for a monodisperse sample can be analyzed via the equation:

$$g^{(2)}(\tau) = B + \beta \exp(-2\Gamma\tau) \tag{2}$$

where B is the baseline of the correlation function at infinite delay, β is the correlation function amplitude at zero delay, and Γ is the decay time. A nonlinear least squares fitting algorithm can be applied to Eq. (2) to retrieve the correlation function decay time Γ. This is exactly what is done in the ASTRA QELS analysis.

From this point, Γ can be converted to the diffusion constant D for the particle via the relation:

$$D = \frac{\Gamma}{q^2} \tag{3}$$

Here, q is the magnitude of the scattering vector, and is given by

$$q = \frac{4\pi n}{\lambda_0}\sin(\theta/2) \tag{4}$$

where n is the solvent index of refraction, λ_0 is the vacuum wavelength of the incident light, and θ is the scattering angle.

Finally, the diffusion constant can be interpreted as the hydrodynamic radius r_h for a diffusing sphere via the Stokes Einstein equation:

$$r_h = \frac{kT}{6\pi\eta D} \tag{5}$$

where k is Boltzmann's constant and η is the solvent viscosity.

The previous equations provide the tools for analyzing a correlation function from a monodisperse sample, but do not address the effects of polydispersity on the correlation function. One of the first attempts to analyze such data was the method of cumulants. First proposed by Koppel[2], the method of cumulants involves expanding Eq. (2) into the various moments of a distribution. In its simplest expression, this expansion turns Eq. (2) into the following:

$$g^{(2)}(\tau) = B + \beta \exp\left(-2\overline{\Gamma}\tau + \kappa_2\tau^2 - \frac{\kappa_3}{3}\tau^3 \ldots\right) \tag{6}$$

1. B. Chu, *Laser Light Scattering: Basic Principles and Practice*, (Academic, Boston, 1991).
2. D.E. Koppel, "Analysis of macromolecular polydispersity in intensity correlation spectroscopy: The method of cumulants," *J. Chem. Phys.* vol. 57, pp. 4814-4820 (1972).

Here, the decay time is now the average for the distribution, while the higher moments correspond to the variance, or width of the distribution (K_2), the skewness of the distribution (K_3) and so on.

In practice, it is usually only possible to determine the first two moments of the expansion in Eq. (6), that is, the average and variance. These are often referred to as the first and second cumulant. In this simplest form, the method of cumulants then boils down to fitting the correlation function to a Gaussian distribution of decay times; only the average and width of the distribution are obtained.

Application of the Method of Cumulants

In the ASTRA software, a variant of Eq. (6) is used to obtain the first and second cumulants in a nonlinear least squares fit of the correlation function. This variation was derived by Frisken[1], and is given by:

$$g^{(2)}(\tau) = B + \beta \exp\left(-2\bar{\Gamma}\tau\right)\left(1 + \frac{\mu_2}{2!}\tau^2 - \frac{\mu_3}{3!}\tau^3 \ldots\right)^2 \qquad (7)$$

Here, the moments μ_n correspond to the K_n terms in Eq. (6), and are the physical moments about the mean $\bar{\Gamma}$. Eq. (7) is inherently more stable than Eq. (6) when fitting at large delay times τ, thus leading to a more robust analysis of the correlation function than has traditionally been obtained from the method of cumulants.

The results obtained from the fit in the QELSBatch cumulant analysis are the first two moments, $\bar{\Gamma}$ and μ_2 in Eq. (7), as well as the baseline B and amplitude β. The baseline and amplitude values are used in the data filtering algorithm to reject QELS correlation functions after the initial cumulants analysis. However, the first two cumulants are the quantities of interest for assessing the polydispersity of the sample.

The first two moments define a Gaussian distribution in decay times, where the first cumulant gives the mean of the distribution, and the square root of the second cumulant gives the standard deviation. In terms of a distribution for sizes, the decay time distribution can be converted to hydrodynamic radius via equations 3 through 5. Since the radius is inversely proportional to the decay time, the distribution in radius is no

1. B.J. Frisken, "Revisiting the method of cumulants for the analysis of dynamic light-scattering data," *Applied Optics*, vol. 40, pp. 4087-4091 (2001).

longer a symmetric Gaussian. This can be seen in Figure G-2.

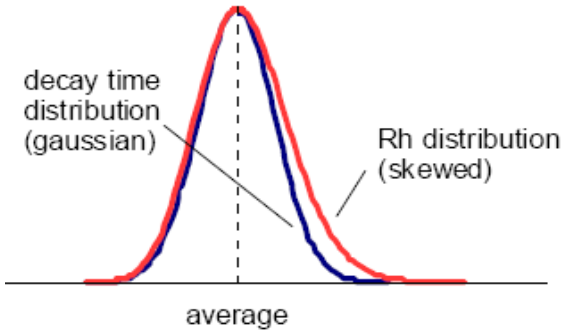


Figure G-2: Cumulants distributions in decay time and hydrodynamic radius

In the cumulants analysis results, the fitted first and second moments—that is the decay time distribution average and variance—are reported, as well as the uncertainties (one standard deviation) from the fit. The square root of the variance is used to determine the standard deviation in the decay time distribution. The average, the average plus the standard deviation, and the average minus the standard deviation are converted to hydrodynamic radius via Eq. 3 through 5, and are included in the results.

Interpretation of Cumulant Results

ASTRA presents the hydrodynamic radius results from the cumulants analysis, since size is more intuitive than decay time for most researchers. In the cumulant analysis graph, the average hydrodynamic radius and the distribution values at one standard deviation are presented. This creates an “error bar” appearance for the graph, but the error bars indicate the width of the fitted distribution. They are asymmetric because the hydrodynamic radius is inversely proportional to the symmetric decay time distribution, as shown in Figure G-2.

With cumulant results presented this way, it is important to remember that there is uncertainty in the first and second moments determined from the fit. This uncertainty translates into an uncertainty in the average hydrodynamic radius, but more importantly, an uncertainty in the effective width of the distribution implied by the cumulants analysis graph.

Therefore, an uncertainty in this width is estimated by calculating the spread in possible width values based on the fitted uncertainty in the second cumulant. The effective width implied by the cumulants is then compared to the spread of possible widths to derive a percentage uncertainty in the effective width. The average uncertainty in width is reported in the Width property, and should provide a good measure of how much to trust the widths that result from the analysis.

After all is said and done, the question remains how the cumulant analysis results relate to the actual polydispersity of the sample. Assuming that the size distribution in Figure G-2 reflects all samples is simply incorrect. Therefore, the cumulant results should be taken as a semi-quantitative estimate of the degree of polydispersity. It would

probably be safe to assume that for two samples with the same average size, but different widths estimated from the cumulant analysis, that the sample with the greater width is more polydisperse. However, trying to define a rigorous polydispersity index from the cumulant analysis would probably lead to very inaccurate results when compared to a quantitative method such as fractionation followed by light scattering to determine the underlying distribution. Therefore, cumulant analysis results should only be used to assess the potential relative polydispersity of samples. Follow-up analysis, such as fractionation followed by light scattering, should be used to assess the reliability of the cumulant analysis results, particularly if they are to be used as the sole assay for polydispersity.

Regularization Theory

Whereas the method of cumulants is one of the simplest approaches to analyzing QELS data from a polydisperse sample, the regularization analysis is one of the most sophisticated. There are many excellent references for the regularization method, and the theory is quite detailed.[1]

Regularization Analysis Histograms

As opposed to the method of cumulants, the regularization analysis makes far fewer assumptions about the underlying distribution of sizes that make up the polydisperse sample. A simple predecessor of the regularization method—the histogram method—demonstrates this nicely. In the histogram method, the distribution of decay times is not assumed to be Gaussian, as it is for the cumulant method with only the first two cumulants. Instead, the decay time distribution is divided into bins. Consider, for example, the model correlation function in Figure G-3 for a bimodal distribution consisting of widely separated sizes:

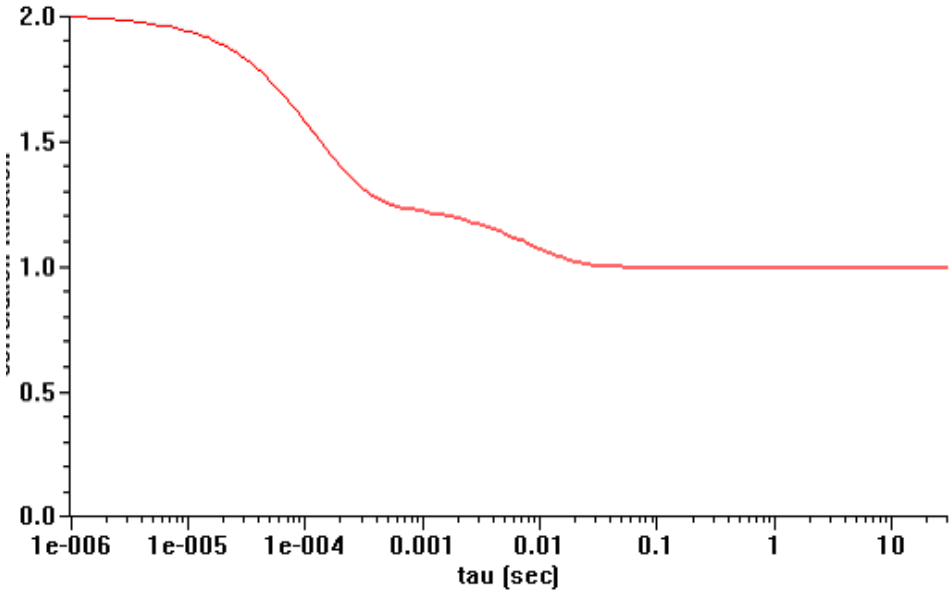


Figure G-3: Bimodal correlation function for mixture of 10 nm and 1 µm particles.

Clearly, the correlation function in Figure G-3 would best be fit by a "sum" of two separate correlation functions, one with a short decay time, and one with a long decay time. In terms of the histogram method, the underlying distribution would appear as in Figure G-4, that is, only two bins would have any intensity. The correlation function is then modeled by "adding"

1. S.W. Provencher, "Inverse problems in polymer characterization: Direct analysis of polydispersity with photon correlation spectroscopy," *Makromol. Chem.*, vol. 180, pp. 201-209 (1979).

the correlation functions for the two separate bins. (It is more complicated than this, since there is cross-correlation between the various components, but for the sake of pedagogy, the concept of adding is adequate.)

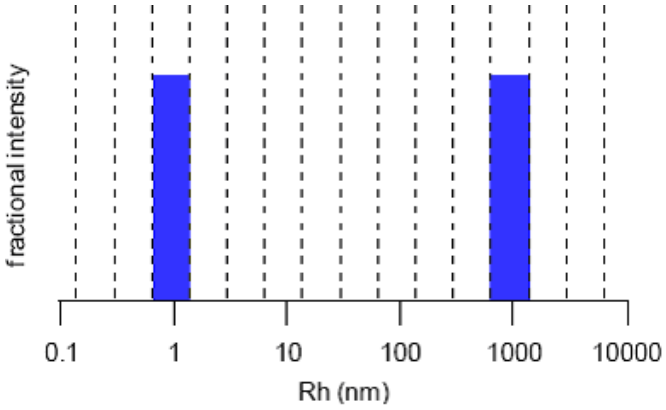


Figure B-2. Model histogram for bimodal size distribution.

Figure G-4: Model histogram for bimodal size distribution

More complicated correlation functions from more polydisperse samples could be modeled by the histogram method. Intensity would be shifted between bins until the right match was found. In so doing, the underlying distribution would be revealed, albeit in a somewhat jagged fashion from the bins. We can remedy this jaggedness by making our bin sizes smaller and smaller until we get the true distribution.

In reality, the histogram method breaks down long before enough bins can be added to accurately represent a distribution. The problem is that as more bins are added, the number of possible solutions explodes. There is not enough information in the correlation function to accurately distribute the intensity over hundreds of bins. In fact, for a standard correlator such as that in the WyattQELS instrument, the largest number of bins that can be handled is only about ten.

The regularization method makes it possible to have a finer mesh of bins. This is accomplished by constraining the types of distributions that can accurately reproduce the correlation function. The most common constraint, and the one employed in the DYNALS algorithm used in ASTRA, is that the distribution be smooth. This is accomplished by adding a regularization term that penalizes solutions that are not smooth. The magnitude of the regularization term determines how smooth the final result must be. The trick of every regularization algorithm is to determine the optimal amount of regularization such that the final solution captures as many features of the true distribution as possible, while balancing out the effects of noise in the correlation function. Noise can add spurious components to the calculated distribution, hence as the noise increases, the regularization term needs to increase to damp these spurious components.

Implementation of Regularization in ASTRA

The regularization algorithm in the ASTRA software is the DYNALS 2.0 algorithm supplied by Alango, Ltd.[1] The DYNALS algorithm sets the regularization level—referred to as the resolution—to the most appropriate value for the level of noise in the correlation function. The resolution value can range between 0 and 1, where 0 corresponds to the noisiest data, and 1 corresponds to the least noisy data. In ASTRA, the optimal value of the resolution is taken from the DYNALS algorithm and reported in the data window for the regularization analysis window.

The results of the regularization are an intensity distribution in hydrodynamic radius. However, in light scattering, the intensity distribution does not give an accurate representation of the number distribution. Therefore, intensity information can be converted to relative number by choosing a mass model for the particles, and applying a correction factor for the intensity. The mass models in ASTRA are sphere and random coil.

Interpreting Regularization Results

Regularization analysis results are more physical than results for the cumulants method. However, some care must be taken in interpreting these results. First, low size peaks (< 1 nm) often appear in the regularization results. These are sometimes attributed to solvent scattering, but are most likely due to avalanche photodiode afterpulsing picked up by the correlator. To exclude this from the correlation function, try setting a longer minimum delay time for the correlation function in ASTRA. Large size peaks are also common in the final distribution. These are usually real and correspond to dust.

Another issue of concern in interpreting regularization results is determining whether the resulting width of the distribution corresponds to an actual polydispersity. For example, applying the regularization analysis to a correlation function from a monodisperse sample often results in a distribution with some width. In general, the noisier the correlation function, the lower the optimal resolution of the regularization algorithm, and the broader the apparent width. Therefore, when interpreting distribution widths from regularization, always consider the resolution obtainable given the level of noise in the correlation function. Ideally, correlation functions for a monodisperse sample and the sample of interest can be obtained with comparable levels of noise, such that the regularization analysis resolution can be accurately assessed.

Finally, the smoothing nature of the regularization algorithm can mask features in the true distribution, even for correlation functions with very low noise. Therefore, if a very structured distribution in sizes is expected, regularization typically returns a much smoother distribution. In short, it

1. A.A. Goldin, “Software for particle size distribution analysis in photon correlation spectroscopy,” website documentation at http://www.softscientific.com/science/WhitePapers/dynals1/dynals100.htm.

is prudent to compare regularization results with a quantitative method such as fractionation followed by light scattering to determine the true distribution. In general, regularization provides the most accurate analysis for samples that are broadly polydisperse over several orders of magnitude in size and that have intrinsically smooth distributions.

H Viscosity Theory

This appendix reviews the theory of viscosity-related calculations.

CONTENTS **PAGE**

Calculating Intrinsic Viscosity 426
Intrinsic Viscosity and Molecular Parameters 428
Flory-Fox Relation 429

Calculating Intrinsic Viscosity

ASTRA 7 can process a wide variety of input viscosity sources, ranging from simple devices producing only a single pressure differential to more sophisticated devices that measure specific viscosity directly.

Once specific viscosity is measured, it is useful to compute the intrinsic viscosity. Intrinsic viscosity is defined as the limit of:

$$[\eta] = \lim_{c \to 0} \frac{\eta_{sp}}{c} \tag{1}$$

Of course, all real instruments measure the specific viscosity at finite concentrations. The concentration dependency of the specific viscosity is typically described using one of three formalisms: the Huggins equation, the Kraemer equation, and the Solomon-Gatesman equation. In all cases, the concentration of the sample must be derived from a detector, such as the Optilab or a UV absorption detector.

Huggins Equation

The Huggins equation is specified as:

$$\eta_{sp} = [\eta]c + k'[\eta]^2 c^2 + O\left(c^3\right) \tag{2}$$

The coefficient k' is the Huggins constant. For random coil polymers in good solvents, the Huggins constant typically has a value between 0.0 and 0.3. In size-exclusion chromatography, the concentration of the sample is usually so dilute that one can ignore the concentration terms of third power (or above) and use the approximation:

$$[\eta] = \eta_{sp} / c \tag{3}$$

Solving Eq. (2) for intrinsic viscosity yields:

$$[\eta] \approx \frac{\eta_{sp}}{c} - \frac{\eta_{sp}^2 k'}{c} + O\left(\eta_{sp}^3\right) \tag{4}$$

Eq. (4) is simplified to the following for computation:

$$[\eta] \approx \frac{-1 + \sqrt{1 + 4\eta_{sp} k'}}{2k'c} \tag{5}$$

Kraemer Equation

The Kraemer equation is:

$$\frac{\ln(\eta_{sp}+1)}{c} \approx [\eta] + k''[\eta]^2 c \tag{6}$$

Solving Eq. (6) for intrinsic viscosity yields:

$$[\eta] \approx \frac{\eta_{sp}}{c} - \frac{(1/2 + k'')\eta_{sp}^2}{c} + O(\eta_{sp}^3) \tag{7}$$

The expansion shows that for small values of specific viscosity, which is almost always the case for chromatography, the two formalisms are related:

$$k' = 1/2 + k'' \tag{8}$$

Eq. (7) is simplified to the following for computation:

$$[\eta] \approx \frac{-1 + \sqrt{1 + 4k''\ln(1+\eta_{sp})}}{2k''c} \tag{9}$$

Solomon-Gatesman Equation

The advantage of the Solomon-Gatesman equation is that it does not require empirical constants. However, for values of specific viscosity much less than one, it reduces to the Huggins Equation, with a value of k' = 1/3.

$$[\eta] \approx \frac{\eta_{sp}}{c} - \frac{\eta_{sp}^2}{3c} + O(\eta_{sp}^3) \tag{10}$$

Eq. (10) is simplified to the following for computation:

$$[\eta] \approx \frac{\sqrt{2\eta_{sp} - 2\ln(\eta_{sp}+1)}}{c} \tag{11}$$

Intrinsic Viscosity and Molecular Parameters

The simplest model of the intrinsic viscosity is due to Einstein and Simha[1]. They considered the case of noninteracting rigid particles. They found that the viscosity can be related to the volume fraction of the fluid occupied by the particles. They found:

$$\eta = \eta_0(1 + \gamma\phi) \tag{12}$$

where ϕ is the volume fraction and $\gamma = 2.5$ for spheres and larger for nonspherical particles.

If the weight concentration of the molecule is c, then the number of molecules per unit volume is $N_A c/M$, where N_A is Avogadro's number and M is the molar mass as measured by light scattering. Therefore Eq. (12) can be written in terms of the measured intrinsic viscosity as:

$$[\eta] = \frac{\gamma N_A V_h}{M} \tag{13}$$

where V_h is the hydrodynamic volume of the molecules. Note that M/V_h is the molecular density, so in some sense, the intrinsic viscosity is measuring the molecular density.

The intrinsic viscosity often differs from the bulk density due to molecular shape, molecular density, and the effects of adsorbed or immobilized solvent on the surface of molecule. This so-called hydration layer moves with the molecule, so it affects measurement of the molecular density. In addition, when the molecule has an extended shape, penetration of non-immobilized solvent into the interior of the molecule similarly affects this measurement.

If we set $\gamma = 2.5$, this can be used to define the equivalent spherical volume of a nonspherical molecule. Similarly, it can be used to define the hydrodynamic volume r_h as:

$$r_h = \left[\frac{3V}{4\pi}\right]^{1/3} \tag{14}$$

When defined in this way, r_h is the radius of a sphere with the same intrinsic viscosity as the molecule under study.

1. A. Einstein, "Eine neue Bestimmung der Molekiildimensionen," *Ann.Physik*, vol. 19, pp. 289-306 (1906).
   A. Einstein, "'Berichtigung zu meiner Arbeit: Eine neue Bestimmung der Molekuldimensionen," *Ann.Physik*, vol. 34, p. 591-592 (1911).
   R.Simha, "The Influence of Brownian Movement on the Viscosity of Solutions," *J.Phys. Chem.*, vol. 44, pp. 25-34 (1940).
   J.W. Mehl, J.L. Oncley, and R. Simha, "Viscosity and the Shape of Protein Molecules," *Science*, vol. 92, pp. 132-133 (1940).

Flory-Fox Relation

While the Einstein-Simha relation can be used to define the hydrodynamic radius for solid molecules with adsorbed solvation layers, it not simply related to the molecular size of extended molecules such as random coil polymers. Several models have been developed to consider the effect the hydrodynamic drag on the intrinsic viscosity.

One of the most successful models comes from Flory and Fox who modeled the random coil as a series of "beads on a string" or a "jointed chain". The string is flexible, but beads are rigid. Flory and Fox considered that hydrodynamic friction causes the solvent near the center of the molecule to move with the same velocity as the center of mass, but solvent near the edges is free to flow into and out of the molecule. This led them to a relationship between the intrinsic viscosity and the mean square radius of the polymer chain in a theta solvent. Their model is:

$$[\eta] = \Phi \left\langle r^2 \right\rangle^{3/2} / M \tag{15}$$

where $\langle r^2 \rangle$ is the mean squared end-to-end distance of the chain, and Φ_0 is a universal constant having the value 2.87×10^{23}. In practice, this constant varies somewhat from polymer to polymer with an experimental value closer to 2.5×10^{23}.

The Flory-Fox relationship is valid for polymers in theta solvents. Ptitsyn and Eizner considered the modification required to model other solvents. They found the following relationship:

$$[\eta] = \Phi(\varepsilon) \left\langle r^2 \right\rangle^{3/2} / M \tag{16}$$

$$\Phi(\varepsilon) = \Phi_0 \left(1 - 2.63\varepsilon + 2.86\varepsilon^2\right) \tag{17}$$

where Φ is now a function of the polymer-solvent iteration parameter ε, and Φ_0 is the Flory-Fox constant. When $\varepsilon = 0$, it reduces to the theta solvent result.

The ε parameter is experimentally measurable with a Mark-Houwink analysis. To perform a Mark-Houwink analysis, the data for a random coil polymer is fit to:

$$[\eta] = KM^a \tag{18}$$

where M is the molar mass. The K and a are fit parameters, which depend upon the polymer, solvent, and temperature. Traditionally, this data is also plotted as Log[η] vs. Log[a]. If the data is fit well, this should be a straight line. The slope parameter a is related to ε by:

$$\varepsilon = (2a - 1)/3 \tag{19}$$