

Analysis of Macromolecular Polydispersity in Intensity Correlation Spectroscopy: The Method of Cumulants

DENNIS E. KOPPEL*

IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598

(Received 3 August 1972)

The first order electric field correlation function of laser light scattered by polydisperse solutions of macromolecules can be written as a sum or distribution of exponentials, with decay rates proportional to the diffusion coefficients of the solute molecules. It is shown that the logarithm of this correlation function is formally equivalent to a cumulant generating function. A method is described by which the distribution function of the decay rates (and thus the extent of polydispersity) can be characterized, in a light scattering experiment, by calculation of the moments or cumulants. The systematic and random statistical errors in the calculated cumulants are discussed.

I. INTRODUCTION

In recent years, the spectral analysis of scattered laser light has proven to be a useful technique for the study of the dynamics of macromolecules in solution.¹ In addition to the conventional total scattered intensity, one can now measure the first-order electric field correlation function,

$$g^{(1)}(|t_2 - t_1|) \equiv \langle \epsilon^*(t_1)\epsilon(t_2) \rangle / \langle |\epsilon|^2 \rangle; \quad (1)$$

the angular brackets indicating an ensemble or infinite time average.

For simplicity, consider the light scattered by dilute solutions of optically isotropic molecules, which are small compared to a wavelength of light, or spherically symmetric, or both.

For a monodisperse solution,²

$$|g^{(1)}(\tau)| = \exp(-\Gamma\tau), \quad (2)$$

where

$$\Gamma = DK^2,$$

with D , the translational diffusion coefficient of the molecules, and K , the magnitude of the scattering vector (the momentum transfer in units of \hbar). For polydisperse solutions, Eq. (2) must be generalized to a sum or distribution of exponentials:

$$|g^{(1)}(\tau)| = \int_0^\infty G(\Gamma) \exp(-\Gamma\tau) d\Gamma,$$

with

$$\int_0^\infty G(\Gamma) d\Gamma = 1. \quad (3)$$

The distribution function of the decay rates, $G(\Gamma)$, can be a broad continuous distribution, a series of discrete delta functions, or some combination of the two. $G(\Gamma)d\Gamma$ is the fraction of the total intensity scattered, on average, by molecules for which $DK^2 = \Gamma$, within $d\Gamma$. Depending upon the aim of a given experiment, polydispersity is something to be studied in itself, merely taken into account, or strictly guarded against. The object of data analysis, in any event, is to charac-

terize $G(\Gamma)$. There are at least three approaches to the problem. First, one can calculate $G(\Gamma)$ directly, inverting the Laplace integral equation [Eq. (3)], with Fourier transforms.³ This technique, however, requires data of high precision over an extraordinary range, and has not been widely employed. Alternately, one may presuppose a specific form for $G(\Gamma)$, and calculate the parameters which give the best fit to the data. Taking this distribution specific approach, several authors⁴⁻⁶ have analyzed the particular case of a Schulz distribution of molecular weights. This approach makes good sense when there is a reasonable basis for supposing that a particular distribution has some correspondence to the sample under study.

We now propose a system of analysis which is more general. It is based on the formalism of the statistical cumulant generating function. With a simple fit of the experimental estimates of $\ln |g^{(1)}(\tau)|$ to a polynomial, we characterize $G(\Gamma)$ with a calculation of its moments or cumulants. While this paper relates to the specific technique of photocount autocorrelation, the cumulants data analysis introduced here can be applied equally well to any experiment which yields a sum or distribution of exponentials. A preliminary version of this approach has appeared elsewhere.⁷

II. MOMENTS AND CUMULANTS

The correlation function of the light scattered by polydisperse solutions lends itself naturally to an analysis in terms of moments or cumulants. We observe, first of all, the exact formal correspondence between the form of the correlation function just proposed [Eq. (3)] and the moment generating function,⁸⁻¹⁰

$$\begin{aligned} \mathfrak{M}(-\tau; \Gamma) &\equiv \langle \exp(-\Gamma\tau) \rangle_{av} \\ &= |g^{(1)}(\tau)|. \end{aligned} \quad (4)$$

$\langle \exp(-\Gamma\tau) \rangle_{av}$ here signifies an average over Γ , weighted by the distribution function $G(\Gamma)$. The moments of the distribution are related to the derivatives of

$\mathfrak{N}(-\tau; \Gamma)$ with respect to $(-\tau)$:

$$\begin{aligned} \mu_m(\Gamma) &\equiv \langle \Gamma^m \rangle_{Av} \\ &= [d^m/d(-\tau)^m] \mathfrak{N}(-\tau; \Gamma) |_{-\tau=0}. \end{aligned} \quad (5)$$

Similarly, one can define the cumulant generating function^{8,9} as the natural logarithm of the moment generating function;

$$\begin{aligned} \mathcal{K}(-\tau; \Gamma) &\equiv \ln \mathfrak{N}(-\tau; \Gamma) \\ &= \ln |g^{(1)}(\tau)|. \end{aligned} \quad (6)$$

The m th cumulant of Γ , $K_m(\Gamma)$, is the coefficient of $(-\tau)^m/m!$ in the MacLaurin expansion of $\mathcal{K}(-\tau; \Gamma)$, which thus takes the form of a power series in τ . That is

$$K_m(\Gamma) \equiv [d^m/d(-\tau)^m] \mathcal{K}(-\tau; \Gamma) |_{-\tau=0} \quad (7)$$

and

$$\mathcal{K}(-\tau; \Gamma) = \sum_{m=1}^{\infty} K_m(\Gamma) \frac{(-\tau)^m}{m!}. \quad (8)$$

For a single exponential correlation function, $\mathcal{K}(-\tau; \Gamma)$ is of course, simply linear in τ . Deviations from "exponentiality" of $|g^{(1)}(\tau)|$ thus appear in Eq. (8) as terms of higher order in τ .

The cumulants can be written explicitly in terms of the moments,⁹

$$\begin{aligned} K_m(\Gamma) &= \sum_{l=1}^m (-1)^{l-1} (l-1)! \sum_{\{a\}} c(l; a_1, \dots, a_m) \\ &\quad \times [\mu_1(\Gamma)]^{a_1} \dots [\mu_m(\Gamma)]^{a_m}, \end{aligned} \quad (9)$$

with

$$c(l; a_1, \dots, a_m) = m! / \prod_{j=1}^m a_j! (j!)^{a_j},$$

and the sum over $\{a\}$ including all sets of nonnegative integers a_1, \dots, a_m for which

$$\sum_{j=1}^m j a_j = m$$

and

$$\sum_{j=1}^m a_j = l. \quad (10)$$

It is clear from Eqs. (9) and (10), that the m th cumulant exists if, and only if, all moments of order $\leq m$ exist. All moments of Γ are necessarily finite, however, since physically, the diffusion coefficient of any molecule is finite.

The cumulants, except the first, are invariant under change of origin, i.e., for any constant c , as a direct consequence of Eqs. (4), (6), and (7),

$$K_{m>1}(\Gamma) = K_{m>1}(\Gamma - c),$$

but

$$K_1(\Gamma) = K_1(\Gamma - c) + c. \quad (11)$$

As a result, choosing $c = \langle \Gamma \rangle_{Av}$, the cumulants can be written directly from Eq. (9), in their simplest and most useful form, as combinations of moments about the mean,

$$\begin{aligned} K_1 &= \langle \Gamma \rangle_{Av} = \int_0^{\infty} \Gamma G(\Gamma) d\Gamma, \\ K_2 &= M_2, \\ K_3 &= M_3, \\ K_4 &= M_4 - 3(M_2)^2, \\ &\vdots \end{aligned} \quad (12)$$

where

$$M_m \equiv \langle (\Gamma - \langle \Gamma \rangle)_{Av}^m \rangle_{Av}.$$

The first cumulant, by Eqs. (2) and (12), is directly proportional to some average diffusion coefficient, $\langle D \rangle_{Av}$. If, as is often the case,¹¹ the intensity of light scattered by macromolecular species i is proportional to the molecular weight m_i , times the weight concentration c_i , i.e.,

$$G(\Gamma) = \sum_i c_i m_i \delta(\Gamma - \Gamma_i) / \sum_i c_i m_i, \quad (13)$$

then $\langle D \rangle_{Av}$ is the so-called z average diffusion coefficient,¹²

$$\begin{aligned} \langle D \rangle_{Av} &= \langle \Gamma \rangle_{Av} / K^2 \\ &= D_z, \\ D_z &\equiv \sum_i c_i m_i D_i / \sum_i c_i m_i. \end{aligned} \quad (14)$$

This is a result of some significance; since it is the z average D which, when combined in the Svedberg equation in the usual way with the routinely measured weight average sedimentation coefficient,¹³ gives a well defined result¹²: the weight average molecular weight. The second cumulant, K_2 , suitably normalized by $(K_1)^2$, is a good measure of the relative width of the distribution. Similarly, K_3 and K_4 are measures of the skewness or asymmetry, and the kurtosis, the peakedness, or flatness of the distribution. It is interesting to note, that for a Gaussian distribution of decay rates, all cumulants of order higher than two are identically zero.

III. SPECTRAL EQUIVALENT

For completeness, now, we present the spectral equivalent of this approach. $S_1(\omega)$ and $S_2(\omega)$, below, are

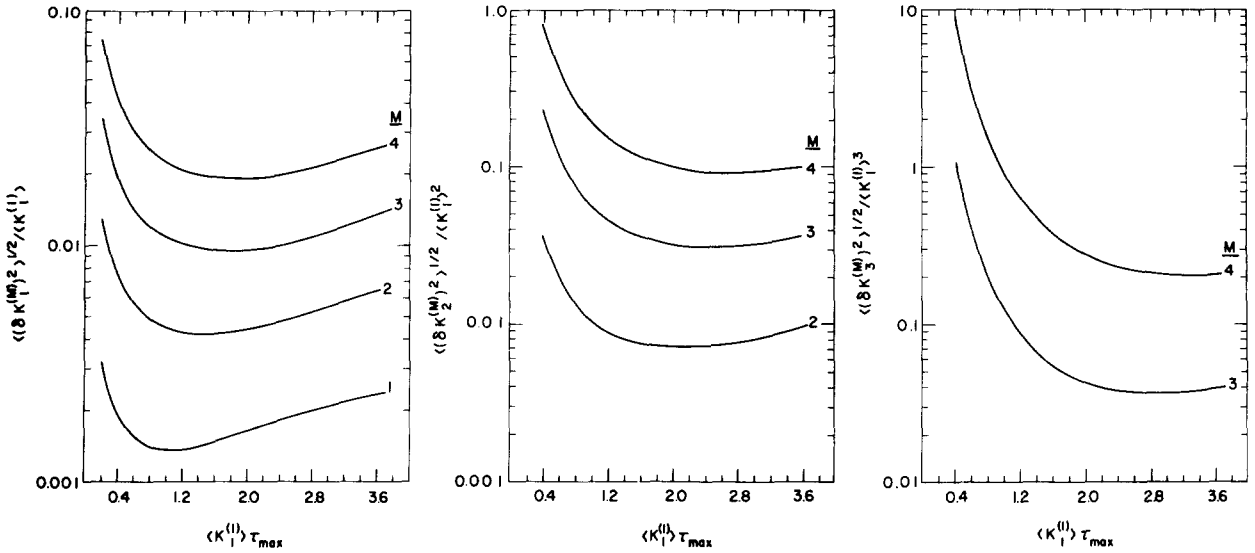


FIG. 1. Normalized standard deviations of the first three cumulants ($m=1-3$), calculated in different order polynomial fits ($M=1-4$), with 20 equally spaced data points, in the extreme, low counting rate, Poisson noise limit, with $\langle B \rangle = 10^6$, as functions of $\langle K_1^{(1)} \rangle \tau_{\max}$. These results are applicable for any $|g^{(1)}(\tau)|$.

the heterodyne and homodyne spectra, respectively,¹

$$\begin{aligned}
 S_1(\omega) &\equiv \frac{A_1}{2\pi} \int_{-\infty}^{\infty} |g^{(1)}(\tau)| \exp(-i\omega\tau) d\tau \\
 &= \frac{A_1}{\pi} \sum_{m=1}^{\infty} (-1)^m M_m \frac{\text{Re}[(\langle \Gamma \rangle_{Av} - i\omega)^{m+1}]}{(\langle \Gamma \rangle_{Av}^2 + \omega^2)^{m+1}} \\
 &= \frac{A_1 \langle \Gamma \rangle_{Av}}{\pi (\langle \Gamma \rangle_{Av}^2 + \omega^2)} \left(1 + \frac{\langle \Gamma \rangle_{Av}^2 - 3\omega^2}{(\langle \Gamma \rangle_{Av}^2 + \omega^2)^2} M_2 \right. \\
 &\quad \left. - \frac{\langle \Gamma \rangle_{Av}^4 - 6\langle \Gamma \rangle_{Av}^2 \omega^2 + \omega^4}{\langle \Gamma \rangle_{Av} (\langle \Gamma \rangle_{Av}^2 + \omega^2)^3} M_3 + \dots \right), \quad (15)
 \end{aligned}$$

$$\begin{aligned}
 S_2(\omega) &\equiv \frac{A_2}{2\pi} \int_{-\infty}^{\infty} |g^{(1)}(\tau)|^2 \exp(-i\omega\tau) d\tau \\
 &= \frac{A_2}{\pi} \sum_{m=0}^{\infty} \sum_{\{m_1+m_2=m\}} \frac{(-1)^{m_1 m_2}}{m_1! m_2!} M_{m_1} M_{m_2} \\
 &\quad \times \frac{\text{Re}[(2\langle \Gamma \rangle_{Av} - i\omega)^{m+1}]}{[2(\langle \Gamma \rangle_{Av}^2 + \omega^2)]^{m+1}}. \quad (16)
 \end{aligned}$$

The normalization factors, A_1 and A_2 , are just the integrated spectral intensities. The higher-order terms are again functions of the moments of Γ about its mean, but their complicated dependence on ω makes one appreciate the simple power series form applicable in the time correlation analysis. Similar expressions have been derived previously¹⁰ for the limit $\omega \gg \Gamma$.

IV. CALCULATING CUMULANTS

In this section, we detail a procedure for calculating the cumulants starting with the correlation data of a

digital homodyne experiment. As background for the signal to noise calculations of the following section, we include somewhat more of the details than would otherwise be necessary.

A full, or unclipped, correlator partitions time into a series of clock intervals and constructs the sums of products,

$$c(\tau) \equiv \sum_{i=1}^N n(t_i) n(t_i + \tau), \quad (17)$$

where $n(t_i)$ is the number of photocounts in the clock interval centered about t_i . We define a normalized "signal"⁷

$$s(\tau) \equiv [c(\tau)/B] - 1, \quad (18)$$

where B , the background or accidental correlation level, is calculated from the single channel counting rates,

$$B = N^{-1} \left[\sum_{i=1}^N n(t_i) \right]^2. \quad (19)$$

If the clock interval is short compared to the field correlation time, then¹

$$\langle s(\tau) \rangle = \beta |g^{(1)}(\tau)|^2, \quad (20)$$

where β , the constant of proportionality, incorporates the effects of incomplete spatial coherence of the field at the photodetector.¹⁴ If we now define

$$k(\tau) \equiv \frac{1}{2} \ln s(\tau), \quad (21)$$

and if

$$\langle [\delta s(\tau)]^2 \rangle \equiv \langle [s(\tau) - \langle s(\tau) \rangle]^2 \rangle \ll \langle s(\tau) \rangle^2,$$

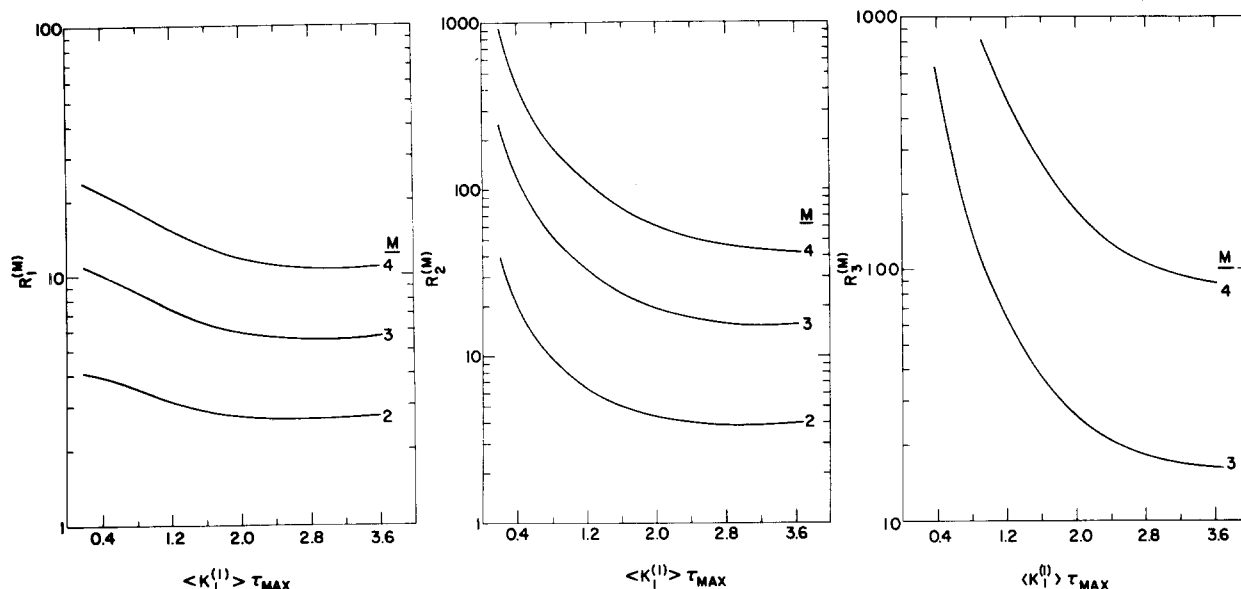


FIG. 2. Normalized cumulant standard deviations relative to that of the single exponential fit,

$$R_m^{(M)} \equiv \frac{\langle (\delta K_m^{(M)})^2 \rangle^{1/2} / \langle K_1^{(1)} \rangle^m}{\langle (\delta K_1^{(1)})^2 \rangle^{1/2} / \langle K_1^{(1)} \rangle},$$

as functions of $\langle K_1^{(1)} \rangle \tau_{MAX}$.

then, by Eqs. (21), (20), (6), and (8),

$$\begin{aligned} \langle \hat{k}(\tau) \rangle &= \frac{1}{2} \ln \langle s(\tau) \rangle \\ &= c_0 + \ln |g^{(1)}(\tau)| \\ &= c_0 + \mathcal{K}(-\tau; \Gamma) \\ &= c_0 + \sum_{m=1}^{\infty} K_m \frac{(-\tau)^m}{m!}, \end{aligned} \quad (22)$$

where c_0 is a constant, independent of $G(\Gamma)$.

The majority of correlators today perform single-clipped correlation,^{7,14-16} in which $n(t_i)$ in the delayed channel is converted into a binary clipped count, $n_q(t_i)$, by the following formula:

$$\begin{aligned} n_q(t_i) &= 0 & \text{if } n(t_i) \leq q \\ &= 1 & \text{if } n(t_i) > q. \end{aligned} \quad (23)$$

The above relationship, between the experimental data and the cumulant generating function [Eq. (22)] holds true for the equivalent clipped signal, defined with

$$\begin{aligned} c(\tau) &\equiv \sum_{i=1}^N n_q(t_i) n(t_i + \tau), \\ B &\equiv N^{-1} \left[\sum_{i=1}^N n_q(t_i) \right] \left[\sum_{i=1}^N n(t_i) \right]. \end{aligned} \quad (24)$$

We calculate the cumulants, for the data of a particular experiment, with a weighted least-squares fit of $\hat{k}(\tau)$

[see Eq. (22)] to the M th-order polynomial,

$$c_0^{(M)} + \sum_{m=1}^M K_m^{(M)} \frac{(-\tau)^m}{m!},$$

with superscripts designating the order of the fit. In the usual fashion,^{7,17} each data point, $\hat{k}(\tau_j)$, is weighted by estimates of $\langle [\delta \hat{k}(\tau_j)]^2 \rangle^{-1}$.

In this notation, then, $K_m^{(M)}$ is the experimental estimate of K_m . There are two sources of error in this estimate. First, there is the obvious effect of the statistical deviations from the average in the measured values of $\hat{k}(\tau_j)$. The statistical errors limit the value of M (there are $M+1$ parameters in the fit) that one can profitably employ.

Also, the forced fitting to a finite polynomial of a function that is, in general, an infinite series results in systematic errors. Therefore, $K_m^{(M)}$ is a biased estimate, in that even the average, $\langle K_m^{(M)} \rangle$, differs from K_m . We evaluate these two effects in the following section.

V. PRECISION AND ACCURACY

A. Random Error

The least squares problem is analytically solvable for a polynomial fit.¹⁷ Each $K_m^{(M)}$, in our case, is a linear combination of the data points, $\hat{k}(\tau_j)$. The variance, $\langle (\delta K_m^{(M)})^2 \rangle$, is therefore a linear combination of terms of the form [see Eq. (21)]

$$\langle \delta \hat{k}(\tau_i) \delta \hat{k}(\tau_j) \rangle \simeq \langle \delta s(\tau_i) \delta s(\tau_j) \rangle / 4 \langle s(\tau_i) \rangle \langle s(\tau_j) \rangle, \quad (25)$$

TABLE I. The normalized covariance, $\langle \delta K_m^{(M)} \delta K_n^{(N)} \rangle / [\langle (\delta K_m^{(M)})^2 \rangle \langle (\delta K_n^{(N)})^2 \rangle]^{1/2}$, calculated with 20 equally spaced data points, with $\langle K_1^{(1)} \rangle_{\tau_{\max}} = 2.0$.

	$K_1^{(1)}$	$K_1^{(2)}$	$K_2^{(2)}$	$K_1^{(3)}$	$K_2^{(3)}$	$K_3^{(3)}$
$K_1^{(1)}$	1.00	0.37	-4.3×10^{-7}	0.17	-4.5×10^{-7}	-4.6×10^{-7}
$K_1^{(2)}$	0.37	1.00	0.93	0.46	0.21	-4.6×10^{-7}
$K_2^{(2)}$	-4.3×10^{-7}	0.93	1.00	0.43	0.22	-1.7×10^{-7}
$K_1^{(3)}$	0.17	0.46	0.43	1.00	0.96	0.89
$K_2^{(3)}$	-4.5×10^{-7}	0.21	0.22	0.96	1.00	0.97
$K_3^{(3)}$	-4.6×10^{-7}	-4.6×10^{-7}	-1.7×10^{-7}	0.89	0.97	1.00

where

$$\delta k(\tau) \equiv k(\tau) - \langle k(\tau) \rangle,$$

$$\delta s(\tau) \equiv s(\tau) - \langle s(\tau) \rangle.$$

For simplicity, we consider first the extreme low counting rate, Poisson noise limit. In this case, the error is uncorrelated between data points, and the unclipped and clipped at zero correlation functions are virtually identical, with¹⁸

$$\langle [\delta k(\tau_j)]^2 \rangle \simeq [1 + |g^{(1)}(\tau_j)|^2] / 4 \langle B \rangle |g^{(1)}(\tau_j)|^4. \quad (26)$$

$\langle B \rangle$, again, is the average background correlation level [see Eqs. (19) and (24)]. To keep the calculation of $\langle (\delta K_m^{(M)})^2 \rangle$ as general as possible, applicable for any distribution of decay rates, we approximate $|g^{(1)}(\tau_j)|$ in the expression for the variance of $k(\tau_j)$ [Eq. (26)] with the best fit single exponential $\exp(-\langle K_1^{(1)} \rangle \tau_j)$. With this approximation, $\langle (\delta K_m^{(M)})^2 \rangle^{1/2} / \langle K_1^{(1)} \rangle^m$ can be calculated as a function of $\langle K_1^{(1)} \rangle_{\tau_{\max}}$ without specifying a particular correlation function. (τ_{\max} is just the largest delay time.) Figure 1 shows the results of such calculations, with 20 equally spaced data points, for the first three cumulants, with $\langle B \rangle$ set to 10^6 counts. For different experimental run times, the statistical errors scale as $\langle B \rangle^{-1/2}$. We really want to know the magnitude of $\langle (\delta K_m^{(M)})^2 \rangle^{1/2}$ relative to $\langle K_1 \rangle^m$. This can be calculated from $\langle (\delta K_m^{(M)})^2 \rangle^{1/2} / \langle K_1^{(1)} \rangle^m$, for a particular distribution of decay rates, if one only knows $\langle K_1^{(1)} \rangle / K_1$ as a function of $\langle K_1^{(1)} \rangle_{\tau_{\max}}$. An example of this procedure will be presented below.

For single exponential correlation functions, other authors¹⁸⁻²⁰ have calculated the equivalent of our $\langle (\delta K_1^{(1)})^2 \rangle^{1/2} / \langle K_1^{(1)} \rangle$ for a variety of different conditions. We can take advantage of their results by observing that while $\langle (\delta K_m^{(M)})^2 \rangle^{1/2} / \langle K_1^{(1)} \rangle^m$ changes markedly as a function of such things as counting rate and number of data points, the ratio

$$R_m^{(M)} \equiv \frac{\langle (\delta K_m^{(M)})^2 \rangle^{1/2} / \langle K_1^{(1)} \rangle^m}{\langle (\delta K_1^{(1)})^2 \rangle^{1/2} / \langle K_1^{(1)} \rangle} \quad (27)$$

remains essentially unchanged. Figure 2 shows calculated values of $R_m^{(M)}$, for the first three cumulants, again as functions of the total range of delay times.

To calculate the variance of a function of two or more cumulants, one must take into account the correlation

between the errors of different cumulants calculated with the same data. Table I presents selected values of the normalized covariance,

$$\langle \delta K_m^{(M)} \delta K_n^{(N)} \rangle / [\langle (\delta K_m^{(M)})^2 \rangle \langle (\delta K_n^{(N)})^2 \rangle]^{1/2},$$

calculated with $\langle K_1^{(1)} \rangle_{\tau_{\max}} = 2.0$. Note that the statistical errors of the cumulants calculated within the same polynomial fit are nearly 100% correlated. This effect decreases the variance of such functions as $K_2^{(M)} / \langle K_1^{(M)} \rangle^2$, but necessarily increases the variance of others.

B. Systematic Error

As indicated above, systematic error is introduced to $K_m^{(M)}$ by terms in the power series expansion of $\langle k(\tau) \rangle$ [Eq. (22)] beyond the $(M+1)$ th term. The extent to which $\langle K_m^{(M)} \rangle$ differs from K_m can be determined only for specific correlation functions. We choose a function next in simplicity to a single exponential, a sum of two exponentials with diffusion coefficients, and weight factors differing by factors of 2, i.e.,

$$|g^{(1)}(\tau)| = \frac{1}{3} [\exp(-1.5K_1\tau) + 2 \exp(-0.75K_1\tau)],$$

$$G(\Gamma) = \frac{1}{3} [\delta(\Gamma - 1.5K_1) + 2\delta(\Gamma - 0.75K_1)]. \quad (28)$$

This distribution could correspond, for example, to an equilibrium mixture of monomers and tetramers or octamers. Noting that

$$K_2 / \langle K_1 \rangle^2 = 0.125,$$

$$K_3 / \langle K_1 \rangle^3 = 0.03125,$$

$$K_4 / \langle K_1 \rangle^4 = -0.0234, \quad (29)$$

we examine the systematic errors of the first three cumulants in Fig. 3. We see that in just those cases (short delay times, high-order fits), where the random errors are largest (see Fig. 1), the systematic errors are smallest. In all cases, one must try to balance these two effects to minimize the combined error.

$K_1^{(1)}$ is the decay rate obtained with a force fit of the data to a single exponential. It has no fundamental significance in this case, since it varies considerably as a function of delay time range, and is typically 6% or 7% lower than the true average decay rate. With the addition of just the quadratic term, however, this

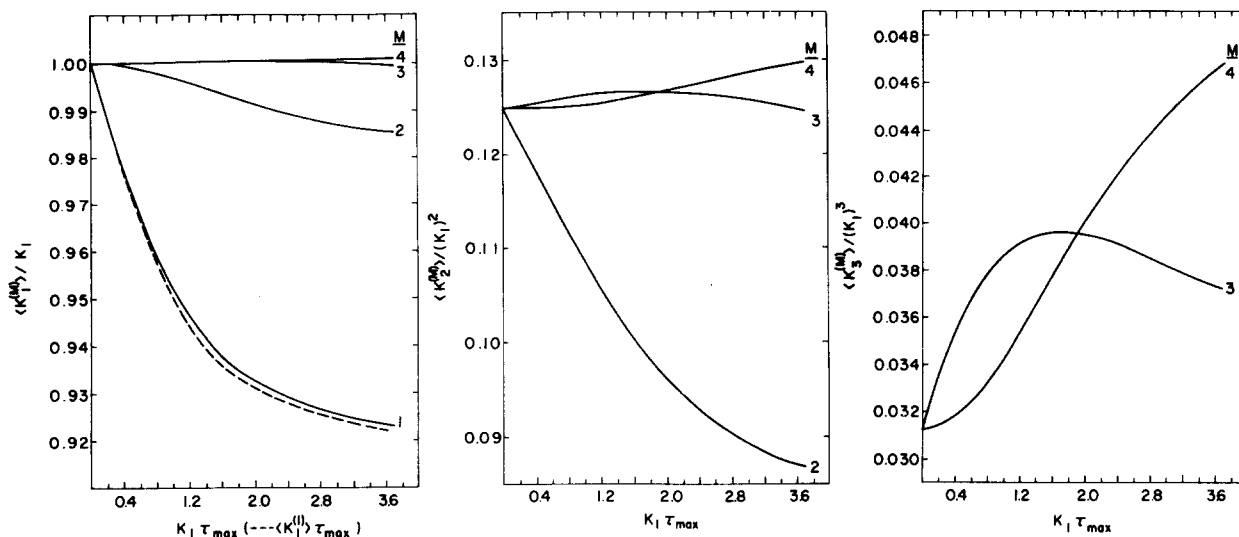


FIG. 3. The first three ($m=1-3$) cumulants for

$$|g^{(1)}(\tau)| = \left\{ \frac{1}{3} [\exp(-1.5K_1\tau) + 2 \exp(-0.75K_1\tau)] \right\},$$

calculated with 20 equally spaced data points, in different order polynomial fits ($M=1-4$) as functions of $K_1\tau_{\max}$.

systematic error is cut to about 1%. The measured value of $K_2^{(2)}$, the coefficient of this first nonexponential correction term, provides a useful quantitative test of monodispersity. It must be zero, within experimental error, before we can safely say that $|g^{(1)}(\tau)|$ is a single exponential.

The dashed curve in Fig. 3 is $\langle K_1^{(1)} \rangle / K_1$ plotted as a function of $\langle K_1^{(1)} \rangle \tau_{\max}$. This is precisely what we need to adapt the general results of Figs. 1 and 2 to the specific correlation function of Eq. (28). For

$$\langle K_1^{(1)} \rangle \tau_{\max} = 2.0,$$

for example, $\langle K_1^{(1)} \rangle / K_1 = 0.931$, so that

$$\begin{aligned} \langle (\delta K_m^{(M)})^2 \rangle^{1/2} / (K_1)^m \\ = (0.931)^m \langle (\delta K_m^{(M)})^2 \rangle^{1/2} / (K_1)^m. \end{aligned} \quad (30)$$

TABLE II. Average cumulants and standard deviations for $|g^{(1)}(\tau)| = \frac{1}{3} [\exp(-1.5K_1\tau) + 2 \exp(-0.75K_1\tau)]$, calculated with 20 equally spaced data points, in the extreme, low counting rate, Poisson noise limit, with $\langle B \rangle = 10^6$ and $\langle K_1^{(1)} \rangle \tau_{\max} = 2.0$.

	M			
	1	2	3	4
$K_1^{(M)} / K_1$	0.931 ± 0.0016	0.991 ± 0.0042	1.000 ± 0.0089	1.000 ± 0.018
$K_2^{(M)} / (K_1)^2$		0.095 ± 0.0062	0.127 ± 0.027	0.127 ± 0.084
$K_3^{(M)} / (K_1)^3$			0.039 ± 0.032	0.041 ± 0.210
$K_4^{(M)} / (K_1)^4$				0.0016 ± 0.210

A similar curve can be constructed for other correlation functions.

C. An Example

We now examine the combined effects of systematic and random errors for a specific example: the two exponential correlation functions of Eq. (28), with $\langle K_1^{(1)} \rangle \tau_{\max} = 2.0$. Figure 4 shows computer-generated data complete with statistical errors consistent with the level of precision assured for Fig. 1 (low counting rate limit, with $\langle B \rangle = 10^6$). We see here the relative insensitivity of light scattering data to solution polydispersity. Systematic deviations from the best straight line fit, the solid line, but can be seen; but these deviations are

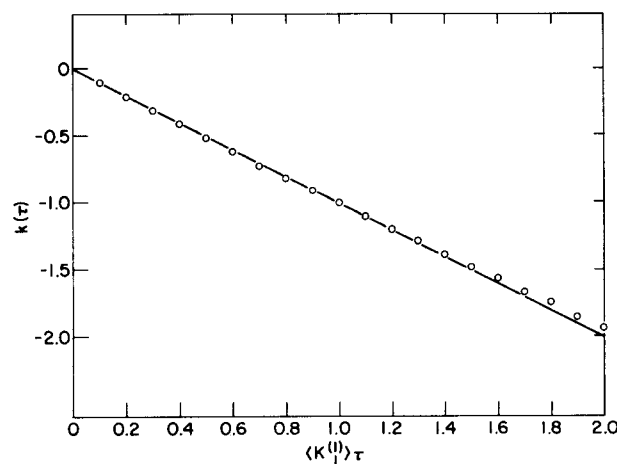


FIG. 4. Sample computer-generated data, with statistical errors consistent with extreme, low counting rate, Poisson noise limit with $\langle B \rangle = 10^6$. $\langle k(\tau) \rangle = \ln \frac{1}{3} [\exp(-1.5K_1\tau) + 2 \exp(-0.75K_1\tau)]$. Solid line is best straight line fit.

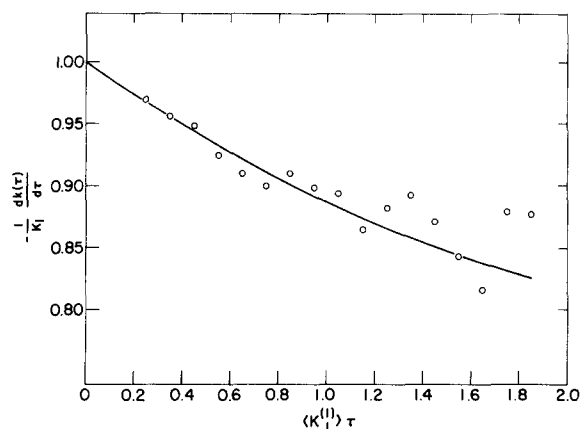


FIG. 5. Numerical derivative of data of Fig. 4. Solid curve is derivative of $\langle k(\tau) \rangle$.

small. As a useful visual display, Fig. 5 shows a calculated derivative of these data.^{7,21} For the derivatives with respect to τ we take the straight-line, least squares slopes of all sets of four consecutive points. The solid line, here, is the derivative of data without errors. The intercept of such a plot, the initial decay rate, is just equal to K_1 , the average decay rate. The general downward slope and upward curvature of the derivative of Fig. 5 are evidence of K_2 and K_3 , respectively. Guided by Eq. (7), we originally employed this numerical differentiation in our calculation of the cumulants.⁷ We have subsequently found the direct polynomial fitting to be superior.

Table II summarizes, for such data, the average calculated cumulants, and the statistical standard deviations about these averages. The average values are taken from Fig. 3 with

$$K_1 \tau_{\max} = \langle K_1^{(1)} \rangle \tau_{\max} / 0.931 \\ = 2.15,$$

and should be compared with the exact cumulants of Eq. (29). The standard deviations are taken from Fig. 1, and adjusted with the aid of Eq. (30). In a third-order fit, for example, we can measure K_1 to better than 1%, K_2 to about 20%, but can determine little more than the sign of K_3 . The data are not precise enough to warrant a fourth-order fit. These results indicate what one can and cannot expect to obtain with this method, and show again the relative insensitivity to polydispersity of light scattering data.

It is left to the reader to derive from Fig. 1, or a combination of Fig. 2 and Refs. 18–20, individual standards of “minimum detectable polydispersity.” Ultimately,

however, one is limited, not by the precision of the data, but by the introduction of systematic errors by such things as laser instability, dust in the sample, or the detection of unshifted laser reflections.⁷

VI. CONCLUSIONS

The analysis of field correlation functions in terms of the cumulants of the decay rate is attractive in theory and applicable in practice with data of reasonable statistical accuracy. This approach provides significant quantitative information on the polydispersity of macromolecular solutions, and is of sufficient simplicity and generality to be incorporated routinely in all data analysis.

ACKNOWLEDGMENTS

The author is greatly indebted to Dr. S. H. Koenig, Dr. D. W. Schaefer, and Dr. P. N. Pusey for discussion and criticism of this work.

* Department of Physics, Columbia University, New York, N.Y. 10027.

¹ See, for example: G. B. Benedek, in *Polarization Matière et Rayonnement, Livre de Jubilé en l'Honneur du Professeur A. Kastler* (Presses Universitaires de France, Paris, 1968); H. Z. Cummins and H. L. Swinney, in *Progress in Optics*, edited by E. Wolf (North-Holland, Amsterdam, 1970), Vol. 8; B. Chu, *Ann. Rev. Phys. Chem.* **21**, 145 (1970); and N. C. Ford, Jr., R. Gabler, and F. E. Karasz, *Advan. Chem.* (to be published).

² R. Pecora, *J. Chem. Phys.* **40**, 1604 (1964).
³ D. G. Gardner, J. C. Gardner, and G. Laush, *J. Chem. Phys.* **31**, 978 (1959).

⁴ Y. Tagami and R. Pecora, *J. Chem. Phys.* **51**, 3293 (1969).
⁵ J. E. Frederick, T. F. Reed, and O. Kramer, *Macromolecules* **4**, 242 (1971).

⁶ N. C. Ford, Jr., R. Gabler, and F. E. Karasz, *Advan. Chem.* (to be published).

⁷ P. N. Pusey, D. E. Koppel, D. W. Schaefer, and R. D. Camerini-Otero, (unpublished).

⁸ M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics* (Charles Griffin, London, 1958), Vol. 1.

⁹ C. D. Cantrell, *Phys. Rev. A* **1**, 672 (1970).
¹⁰ D. W. Schaefer, G. B. Benedek, P. Schofield, and E. Bradford, *J. Chem. Phys.* **55**, 3884 (1971).

¹¹ C. Tanford, *Physical Chemistry of Macromolecules* (Wiley, New York, 1967), Chap. 5.

¹² P. Kinell, *Arkiv Kemi* **14**, 327 (1959).
¹³ C. Tanford, Ref. 11, p. 380.

¹⁴ D. E. Koppel, *J. Appl. Phys.* **42**, 3216 (1971).
¹⁵ E. Jakeman and E. R. Pike, *J. Phys. A* **2**, 411 (1969).

¹⁶ R. Foord, E. Jakeman, C. J. Oliver, E. R. Pike, R. J. Blagrove, and E. Peacocke, *Nature* **227**, 242 (1970).

¹⁷ P. G. Guest, *Numerical Methods of Curve Fitting* (Cambridge U. P., London, 1961).

¹⁸ E. Jakeman, E. R. Pike, and S. Swain, *J. Phys. A* **4**, 517 (1971).

¹⁹ V. Degiorgio and J. B. Lastovka, *Phys. Rev. A* **4**, 2033 (1972).

²⁰ E. R. Pike, *J. Phys. Radium* **33**, C1-177 (1972).
²¹ P. N. Pusey, D. W. Schaefer, D. E. Koppel, R. D. Camerini-Otero, and R. M. Franklin, *J. Phys. Radium* **33**, C1-163 (1972).