Influence of statistical errors on size distributions obtained from dynamic light scattering data. Experimental limitations in size distribution determination

1990 J. Phys. A: Math. Gen. 23 1363

(http://iopscience.iop.org/0305-4470/23/8/012)

J. Phys. A: Math. Gen. 23 (1990) 1363-1383. Printed in the UK

Influence of statistical errors on size distributions obtained from dynamic light scattering data. Experimental limitations in size distribution determination

Zbigniew Kojro

Max-Planck-Institut für Biophysik, Kennedy Allee 70, Frankfurt-am-Main 70, Federal Republic of Germany†, and Institute of Experimental Physics, University of Gdansk, Poland

Received 14 October 1988, in final form 18 July 1989

Abstract. In this paper, the relationship between the level of statistical errors in a measured intensity autocorrelation function and the relative size distribution width uncertainty is estimated. The results of the Fourier analysis of statistical errors existing in measured autocorrelation functions indicate that the individual error course can be similar to the systematic distortion. This has been also confirmed experimentally. The problem of the baseline error is also discussed. It is shown that the baseline error is compensated in dynamic light scattering data during normalisation.

1. Introduction

The size distribution of suspended particles has recently been a subject of great interest. Several authors have used the dynamic light scattering method [1-5] for intensity autocorrelation measurements, and then solved a Fredholm integral equation of the first kind [6] in order to obtain a size distribution. In general, the determination of size distributions from a Fredholm integral equation is an 'ill-posed' problem and leads to difficulties because of the existence of an infinite number of oscillating solutions satisfying the investigated integral within experimental errors. Recently, a number of numerical methods have appeared [7-10] which can be employed in solving 'ill-posed' problems.

Experimental work has been undertaken in which size distributions of suspended lipid vesicles and size distributions of suspended standard spheres (Dow-latex type, see [11]) were determined from dynamic light scattering data by employing Provencher's regularisation method [12] (numerical program CONTIN [7]). This method has shown that the final relative error of the size distribution width is much larger (even a few orders of magnitude) than the relative error of the mean radius r. In practice, it is very difficult to obtain reproducible results for the width of the size distribution for the same measured sample.

There are a few possible explanations of this phenomenon: (a) the regularisation method (CONTIN program) is very sensitive for noise configuration in dynamic light scattering data; (b) the baseline error has a significant effect on the final result; (c) samples can be unstable during measurements; (d) there exists another, statistical reason for such large uncertainties in size distribution width determination.

† Address for correspondence.

0305-4470/90/081363+21$03.50

Explanation (c) can be excluded by checking the stability of all important physical parameters and by using standard particles in experiments. On the other hand, the regularisation method applied to simulated autocorrelation functions, with uncorrelated noise added (without baseline errors) [33], gives quite good results.

In general, there exist many sources of distortions of the final result which can be placed in three groups: (x) distortions arising during numerical calculations; (y) 'outer' distortions induced by an imperfection of measuring instruments or by impurities and instabilities in a measured sample; (z) distortions connected with the stochastic nature of detection and diffusion processes.

In this paper, the relationship between the level of statistical fluctuations in a measured autocorrelation function and the size distribution width uncertainty is estimated. Several assumptions, similar to those described by Saleh and Cardoso [13], are used: (i) the sample time interval is much shorter than the field's coherence time; (ii) the optical field is stationary, cross-spectrally pure and Gaussian; (iii) a measured autocorrelation function contains only statistical errors of the (z) type, i.e. we deal with an ideal sample, instrument and calculating method. The fact that the choice of the delay timescale, in intensity correlation measurements, can make an influence on the final result has also been taken into account.

2. Intensity correlation functions in size distribution determination

A second-order autocorrelation function, AF(2), that is available from limited-time dynamic light scattering measurements may be expressed, in agreement with Saleh and Cardoso [13], in the form

$$\begin{aligned} &\hat{G}_{0\iota} = \sum_{m=1}^{N} n(m)n(m+\iota) = \hat{g}_{0\iota}\bar{n}^2 N = (g_\iota + \varepsilon_\iota)\bar{n}^2 N \\ &E\{\hat{g}_{0\iota}\} = g_\iota = 1 + \chi_\iota^2 \qquad E\{\varepsilon_\iota\} = 0 \\ &\hat{n} = (1/N)\sum_{m=1}^{N} n(m) \qquad E\{\hat{n}\} = \bar{n} \end{aligned} \tag{1}$$

where E is the expectation value, $\hat{G}_{0\iota}$, $\hat{g}_{0\iota}$ are 'unbiased' estimators of AF(2) [13, 14], $\hat{n}$ the unbiased estimator of $\bar{n}$ (mean number of counts per sample time), ε_ι the statistical error, N the total number of samples and χ_ι the field autocorrelation function [13]. Expressions (1) are equivalent to

$$\begin{aligned} &\hat{G}_0^{(2)}(\tau) = \bar{B}\hat{g}_0^{(2)}(\tau) = \bar{B}(g^{(2)}(\tau) + \varepsilon(\tau)) \\ &E\{\varepsilon(\tau)\} = 0 \qquad \bar{B} = \bar{n}^2 N \qquad (\tau \text{ is delay time}) \end{aligned} \tag{2}$$

provided that the ratio sample-time to coherence-time, T_s/T_c, is much less than 1 (see [1, 16–18], and appendix B). Here $\bar{B}$ is the expected value of the baseline, $\varepsilon(\tau)$ a function which describes a course of statistical errors in one data set and $g^{(2)}(\tau)$ is the expected intensity autocorrelation function. In figure 3 (section 6) are shown experimental examples of error courses $\varepsilon(\tau)$ (see also sections 3 and 6).

A relative size distribution function, SD, can be determined from the field autocorrelation function $g^{(1)}(\tau)$ which is available from $g^{(2)}(\tau)$ after subtracting the baseline and taking a square root in accordance with the Siegert relation [19–21]

$$g^{(2)}(\tau) = 1 + |g^{(1)}(\tau)|^2 \tag{3}$$

and then by solving a Fredholm integral equation [6]

$$|g^{(1)}(\tau)| = \zeta^{-1} \int_0^\infty n(r) m^2(r) f(r, q) \exp(-D(r) q^2 |\tau|)\, dr \tag{4}$$

where $n(r)$ is the size distribution, $m(r)$ the mass of investigated particles, $f(r, q)$ the form factor, $D(r)$ the diffusion coefficient, q the magnitude of a scattering vector, r the radius of a particle and ζ the normalisation factor. In the case of the measured autocorrelations, one obtains from (3) and (2) by using the measured baseline $\hat{B}$

$$\begin{aligned} &\hat{G}_0^{(1)}(\tau) = (\hat{G}_0^{(2)}(\tau) - \hat{B})^{1/2} = (\bar{B})^{1/2}(|g^{(1)}(\tau)|^2 + \varepsilon(\tau) - \delta B)^{1/2} \\ &\delta B = (\hat{B} - \bar{B})/\bar{B} \qquad \hat{B} = \hat{n}^2 N \end{aligned} \tag{5}$$

or normalised (biased) first-order autocorrelation function, AF(1),

$$\begin{aligned} \hat{g}^{(1)}(\tau) &= ((\hat{G}_0^{(2)}(\tau) - \hat{B})/\hat{B})^{1/2} = (\hat{g}^{(2)}(\tau) - 1)^{1/2} \\ &= (\bar{B}/\hat{B})^{1/2}(|g^{(1)}(\tau)|^2 + \varepsilon(\tau) - \delta B)^{1/2} \\ &\approx (|g^{(1)}(\tau)|^2 + \varepsilon(\tau) - \delta B - \delta B |g^{(1)}(\tau)|^2)^{1/2} \end{aligned} \tag{6}$$

where $\hat{g}^{(2)}(\tau)$ is the biased AF(2) estimator [13, 14]. Both functions, unnormalised (5) and normalised (6), can be used in a relative SD determination because they lead to the same size distribution $n(r)$, distorted by statistical errors, with the accuracy of the factor ((4), (5), (6)). Equation (6) also expresses the fact that the normalisation contribution $\delta B |g^{(1)}(\tau)|^2$ corresponds to the factor $\bar{B}/\hat{B}$ which is insignificant in size distribution or linewidth determination (4).

3. Statistical errors in intensity autocorrelation function

Three basic papers [13, 14, 24] concerning the problem of statistical errors in AF(2) estimators are not directly comparable. Degiorgio and Lastovka [22] have determined the variance for statistical errors in unbiased AF(2) estimators (1), (2) but Jakeman *et al* [14] and Saleh and Cardoso [13] investigated biased AF(2) estimators (6) for the same purpose. However, by expanding exponentials in [14] and [24] in terms of $\gamma = T_s/T_c \ll 1$ (we must remember that the correlation time is defined differently in papers [14] and [22]), by taking into account only unbiased parts in [13] and [14], and assuming that $\tau = \tau_\kappa = \tau_\iota$ in [13], we obtain the common expression for the variance of statistical errors in unbiased AF(2) estimators (2), in the monodisperse (Lorentzian) case:

$$\begin{aligned} \mathrm{Var}(\hat{g}_0^{(2)}(\tau)) = N^{-1}[&(T_c/T_s)(\tfrac{9}{2} + 4(3+x)\,e^{-x} + 3(\tfrac{1}{2}+x)\,e^{-2x}) \\ &+ (2/\bar{n})(3 + 6\,e^{-x} + 3\,e^{-2x}) + (1/\bar{n})^2(1 + e^{-x})] \\ x = 2\Gamma\tau \qquad &\Gamma = 1/T_c. \end{aligned} \tag{7}$$

in the notation of Jakeman *et al* [14]†.

† Note that terms $b_{\iota\kappa}$ in [13] should be completed to

$$b_{\iota\kappa} = 2[1 + \chi_\iota^2 + \chi_\kappa^2 + \chi_{\iota-\kappa}^2 + 2\,\mathrm{Re}(\chi_\iota^* \chi_\kappa \chi_{\iota-\kappa})] + 2[1 + \chi_\iota^2 + \chi_\kappa^2 + \chi_{\iota+\kappa}^2 + 2\,\mathrm{Re}(\chi_\iota \chi_\kappa \chi_{\iota+\kappa}^*)].$$

When $\tau_\iota \neq \tau_\kappa$, we obtain the expression for correlations in $\varepsilon(\tau)$ for unbiased AF(2) estimators [13] ($T_c = W^{-1}$; new arrangement is used compared with the original paper)

$$\begin{aligned}\Lambda_{\iota\kappa}^{(0)} &= \operatorname{cov}\{\hat{g}_{0\iota}, \hat{g}_{0\kappa}\} = E\{\hat{g}_{0\iota}\hat{g}_{0\kappa}\} - E\{\hat{g}_{0\iota}\}E\{\hat{g}_{0\kappa}\} \\ &= \Lambda + \Lambda_\iota + \Lambda_\kappa + \Lambda_{\iota\kappa,r} \\ &= 4N^{-1}\{T_c/T_s + 1/\bar{n}\} + 4N^{-1}\{(T_c/T_s)\operatorname{Re}(\chi_\iota^* y_\iota) + (1/\bar{n})\chi_\iota^2\} \\ &\quad + 4N^{-1}\{(T_c/T_s)\operatorname{Re}(\chi_\kappa^* y_\kappa) + (1/\bar{n})\chi_\kappa^2\} \\ &\quad + N^{-1}\{(T_c/T_s)[z_{\iota-\kappa} + z_{\iota+\kappa} + 2\operatorname{Re}(\chi_\iota^*\chi_\kappa y_{\iota-\kappa}) + 2\operatorname{Re}(\chi_\iota\chi_\kappa y_{\iota+\kappa}^*) + u_{\iota\kappa}] \\ &\quad + (1/\bar{n})[2\chi_{\iota-\kappa}^2 + 2\chi_{\iota+\kappa}^2 + 4\operatorname{Re}(\chi_\iota^*\chi_\kappa\chi_{\iota-\kappa}) + 4\operatorname{Re}(\chi_\iota\chi_\kappa\chi_{\iota+\kappa}^*) + v_{\iota\kappa}]\} \\ v_{\iota\kappa} &= (1/\bar{n})\delta_{\iota\kappa}(1+\chi_\iota^2). \end{aligned} \tag{8}$$

Here the Saleh and Cardoso notation is used and terms χ, y, z, u are defined in appendix A. In the general expression $E\{\hat{g}_{0\iota}\hat{g}_{0\kappa}\}$ (8) there are three types of terms: (i) constant values; (ii) 'single-point correlations' depending on the single position x_ι or x_κ (figure 3); (iii) 'two-point correlations' depending on both positions x_ι and x_κ. All these terms of order N^0 are subtracted from $E\{\hat{g}_{0\iota}\hat{g}_{0\kappa}\}$ when $\operatorname{cov}\{\hat{g}_{0\iota}, \hat{g}_{0\kappa}\}$ is calculated, and in the expression (8) there remain only terms of order N^{-1} which are grouped according to (i)-(iii) (terms of order N^{-2}, such as the variance for bias term, are neglected). The constant term of order N^{-1} is the value for the variance of the baseline error δB (equation (A5) in [14])

$$\operatorname{Var}(\delta B) = E\{\delta B^2\} \approx (2/\bar{n})^2 \operatorname{Var}(\hat{n}) \approx 4N^{-1}(T_c/T_s + 1/\bar{n}) = \Lambda. \tag{9}$$

On the other hand, the same constant value exists in $\Lambda_{\iota\kappa}^{(1)}$ calculated by Saleh and Cardoso [13]†:

$$\Lambda_{\iota\kappa}^{(1)} = -\Lambda + \Lambda\chi_\iota^2\chi_\kappa^2 - \Lambda_\iota(1+\chi_\kappa^2) - \Lambda_\kappa(1+\chi_\iota^2). \tag{10}$$

This constant term disappears from the covariance calculated for biased AF(2) estimators (6)

$$\begin{aligned}\Lambda_{\iota\kappa} &= \operatorname{cov}\{\hat{g}_\iota, \hat{g}_\kappa\} = \Lambda_{\iota\kappa}^{(0)} + \Lambda_{\iota\kappa}^{(1)} \\ &= \Lambda\chi_\iota^2\chi_\kappa^2 - \Lambda_\iota\chi_\kappa^2 - \Lambda_\kappa\chi_\iota^2 + \Lambda_{\iota\kappa,r}. \end{aligned} \tag{11}$$

From (8)-(11) one can conclude that the baseline error is eliminated from dynamic light scattering data after normalisation. In order to prove this, we can employ a kind of formal calculus, which can be helpful in the interpretation of terms existing in covariances (8), (10), (11) calculated by Saleh and Cardoso [13]. First, we should point out that for a limited-time dynamic light scattering experiment the contents of all correlator channels (unbiased AF(2)) can indicate the common shift, compared with the mean value $\bar{n}^2 N$, because the same information about fluctuations $n(m)$ is introduced simultaneously into all memory channels of a correlator through the input counter and multipliers (figure 10 in [23]). The statistical error course $\varepsilon(\tau)$ (in equations (2), (5) and (6)) can be expressed then as a sum of two random variables: constant shift δB_ε plus reduced noise fluctuations $\varepsilon_r(\tau)$ (figure 3), which are mutually uncorrelated in the stationary range of $\varepsilon_r(\tau)$ (random process $\varepsilon_r(\tau)$ is non-stationary when τ is not much larger than T_c; see section 5.2 and appendix A):

$$\varepsilon(\tau) = \varepsilon_r(\tau) + \delta B_\varepsilon \tag{12}$$

† Note that terms $d_{\iota\kappa}$ in [13] should be corrected to

$$d_{\iota\kappa} = -4[1 + (1+\chi_\iota^2)\operatorname{Re}(\chi_\kappa^* y_\kappa) + (1+\chi_\kappa^2)\operatorname{Re}(\chi_\iota^* y_\iota)] + 4\chi_\iota^2\chi_\kappa^2.$$

where $E\{\varepsilon_r(\tau_\iota)\varepsilon_r(\tau_\kappa)\} \to 0$ if $\tau_\iota \to \infty$ or $\tau_\kappa \to \infty$ and $E\{\delta B_\varepsilon \varepsilon_r(\tau)\} \neq 0$ in the non-stationary range of $\varepsilon_r(\tau)$ (figure 3). Next, we can express the biased estimator of AF(2), by using (1), (12) and the definition of δB given in (5), as

$$\begin{aligned}\hat{g}_\iota &= \hat{G}_{0\iota}/\hat{B} \approx \hat{g}_{0\iota}(1-\delta B+\delta B^2)\\ &= g_\iota + \varepsilon_r(\tau_\iota) + \delta B_\varepsilon - \delta B - \delta B\chi_\iota^2 - \delta B\delta B_\varepsilon - \delta B\varepsilon_r(\tau_\iota) + \delta B^2 + \delta B^2\chi_\iota^2\\ E\{\delta B\} &\approx E\{\delta B_\varepsilon\} = E\{\varepsilon_r(\tau_\iota)\} = 0 \qquad (13)\end{aligned}$$

and calculate the covariance matrix $\Lambda_{\iota\kappa}$, in which should appear auto- and cross-correlations for all random variables from (13). We obtain (terms less than those of order N^{-1} are neglected):

$$\begin{aligned}\Lambda_{\iota\kappa} = {} & E\{(\delta B_\varepsilon - \delta B)^2\} + E\{\delta B^2\}\chi_\iota^2\chi_\kappa^2 - E\{B\varepsilon_r(\tau_\iota)\}\chi_\kappa^2 - E\{\delta B\varepsilon_r(\tau_\kappa)\}\chi_\iota^2\\ & + E\{\varepsilon_r(\tau_\iota)\varepsilon_r(\tau_\kappa)\} + E\{\delta B^2 - \delta B_\varepsilon\delta B\}(\chi_\iota^2+\chi_\kappa^2)\\ & + E\{(\delta B_\varepsilon - \delta B)(\varepsilon_r(\tau_\iota)+\varepsilon_r(\tau_\kappa))\} \qquad (14)\end{aligned}$$

where

$$\Lambda_{\iota\kappa}^{(0)} = E\{\delta B_\varepsilon^2\} + E\{\delta B_\varepsilon\varepsilon_r(\tau_\iota)\} + E\{\delta B_\varepsilon\varepsilon_r(\tau_\kappa)\} + E\{\varepsilon_r(\tau_\iota)\varepsilon_r(\tau_\kappa)\} \qquad (15)$$

and

$$\begin{aligned}\Lambda_{\iota\kappa}^{(1)} = {} & -E\{2\delta B_\varepsilon\delta B - \delta B^2\} + E\{\delta B^2\}\chi_\iota^2\chi_\kappa^2 - E\{\delta B\varepsilon_r(\tau_\iota)\}(1+\chi_\kappa^2)\\ & - E\{\delta B\varepsilon_r(\tau_\kappa)\}(1+\chi_\iota^2) + E\{\delta B^2 - \delta B_\varepsilon\delta B\}(\chi_\iota^2+\chi_\kappa^2). \qquad (16)\end{aligned}$$

Comparing (11) and (14) we conclude that there will be no constant term in $\Lambda_{\iota\kappa}$ when†

$$\delta B_\varepsilon = \delta B \qquad (17)$$

and then (8), (10), (11) and (15), (16), (14) are equivalent, respectively, with $\Lambda = E\{\delta B^2\}$, $\Lambda_\iota = E\{\delta B\varepsilon_r(\tau_\iota)\}$, $\Lambda_\kappa = E\{\delta B\varepsilon_r(\tau_\kappa)\}$, $\Lambda_{\iota\kappa,r} = E\{\varepsilon_r(\tau_\iota)\varepsilon_r(\tau_\kappa)\}$. We can say then that the constant shift in $\varepsilon(\tau)$, δB_ε (constant shift in all correlator channels), and the baseline error δB in $\hat{B}$ vary covariantly during the measurement (see also figures 3 and 4, section 6) and that the baseline error δB disappears, with the accuracy of order N^{-1}, from dynamic light scattering data after normalisation, i.e. after subtraction of the measured baseline $\hat{B}$ from unbiased AF(2) estimator (equations (5), (12), (17); see also appendix B). This surprising result we call 'self-compensation in correlation measurements'. In fact, we could expect such a result because in the channel contents, as well as in $\hat{B}$, is stored the same information about the fluctuation of the mean number of counts. This expectation is expressed by Jakeman *et al* ([14] p 521) and simulations made by Oliver [17] indicate that the smallest errors in linewidth determination are achieved when the measured (not expected) baseline is used in the normalisation procedure. In figure 4 (see section 6) we see that shifts δB_ε are compensated after normalisation. When (17) holds then (14) simplifies to (see also (11)):

$$\Lambda_{\iota\kappa} = E\{\delta B^2\}\chi_\iota^2\chi_\kappa^2 - E\{\delta B\varepsilon_r(\tau_\iota)\}\chi_\kappa^2 - E\{\delta B\varepsilon_r(\tau_\kappa)\}\chi_\iota^2 + E\{\varepsilon_r(\tau_\iota)\varepsilon_r(\tau_\kappa)\}. \qquad (18)$$

Here the last term is responsible for correlations in reduced noise fluctuations. This is the only important error term in size distribution or linewidth determination for the following reasons.

† In fact, $\delta B = \delta B_\varepsilon + \frac{1}{4}\delta B_\varepsilon^2 \approx \delta B_\varepsilon$ [35].

(a) The term $E\{\delta B^2\}\chi_\iota^2\chi_\kappa^2$ describes correlation between normalisation contributions $\delta B\chi_\iota^2$ and $\delta B\chi_\kappa^2$ (appearing during division of unbiased AF(2) estimators $\hat{G}_{0\iota}$ and $\hat{G}_{0\kappa}$ by $\hat{B}$; (13)) which are insignificant in size distribution or linewidth determination (see (6), (5) and (4)). This term, however, is important when the comparison between biased AF(2) estimators, $\hat{g}_\iota$, and expected value, g_ι, is performed and then (18) expresses the variance (when $\tau_\iota = \tau_\kappa$) for $\hat{g}_\iota - g_\iota$ (Jakeman *et al* [14], Saleh and Cardoso [13]).

(b) The terms $-E\{\delta B\varepsilon_r(\tau_\iota)\}\chi_\kappa^2$ and $-E\{\delta B\varepsilon_r(\tau_\kappa)\}\chi_\iota^2$ describe the fact that reduced noise fluctuations are correlated with the baseline error (in the non-stationary range of reduced noise fluctuations—see sections 6 and 8) but this cannot have any influence on size distribution or linewidth determination. These terms, however, are important when the variance for $\hat{g}_\iota - g_\iota$ ($\tau_\iota = \tau_\kappa$) is calculated because reduced noise fluctuations, in the non-stationary range, have on average opposite signs than normalisation contributions (see section 8).

(c) Baseline error is compensated during normalisation.

Finally, the following expression is responsible for noise correlations and the variance ($\tau_\iota = \tau_\kappa$) of statistical errors in a measured AF(2), when size distribution or linewidth is determined:

$$\Lambda_{\iota\kappa,r} = E\{\varepsilon_r(\tau_\iota)\varepsilon_r(\tau_\kappa)\} \tag{19}$$

in agreement with (12), (17), (5), (6) and (4). Hence, we have in the monodisperse case

$$\begin{aligned}\Lambda_{\iota\kappa,r} = N^{-1}\{(T_c/T_s)[(\tfrac{1}{2}+|x_\iota - x_\kappa|)\exp(-2|x_\iota - x_\kappa|) + 3(\tfrac{1}{2}+(x_\iota + x_\kappa))\exp(-2(x_\iota + x_\kappa))\\ + 4(1+|x_\iota - x_\kappa|)\exp(-(x_\iota + x_\kappa) - |x_\iota - x_\kappa|)] + (2/\bar{n})[\exp(-2|x_\iota - x_\kappa|)\\ + 2\exp(-(x_\iota + x_\kappa) - |x_\iota - x_\kappa|) + 3\exp(-2(x_\iota + x_\kappa))]\\ + \delta_{\iota\kappa}/(\bar{n}^2)[1+\exp(-2x_\iota)]\} \qquad x_\iota = \Gamma\tau_\iota\end{aligned} \tag{20}$$

and

$$\begin{aligned}\Lambda_{\kappa\kappa,r} = N^{-1}[(T_c/T_s)(\tfrac{1}{2}+4\,e^{-x}+3(\tfrac{1}{2}+x)\,e^{-2x}) + (2/\bar{n})(1+2\,e^{-x}+3\,e^{-2x})\\ + (1/\bar{n})^2(1+e^{-x})] \qquad x = 2\Gamma\tau_\kappa.\end{aligned} \tag{21}$$

This equation agrees almost perfectly with the variance calculated by Schätzel [15] for so-called 'modified' AF(2) estimators (equation (25) in [15]), if the limit $T_s/T_c \to 0$ is considered.

4. The choice of the delay timescale

The problem of the choice of the timescale in correlation measurements has been discussed by several authors [14, 22, 23, 34]. Here, only one aspect of this problem is dealt with, namely distortions which can appear when a square root from AF(2) is calculated in order to obtain AF(1). Based on (6), (12), (17) we can plot the functions (figure 1)

$$\begin{aligned}\Delta g_{\pm\sigma}^{(1)}(\tau) &= |g^{(1)}(\tau)| - (|g^{(1)}(\tau)|^2 \pm \sigma(\tau))^{1/2}\\ \Delta g_{\pm\sigma}^{(2)}(\tau) &= \pm\sigma(\tau)\end{aligned} \tag{22}$$

where the factor $\sqrt{\bar{B}/\hat{B}}$, which has no influence on SD (or linewidth) determination, is replaced by 1 and $\sigma(\tau)$ is a standard deviation for reduced noise fluctuations $\varepsilon_r(\tau)$ in AF(2) (21). Expressions (22) describe limits (standard deviations) for statistical

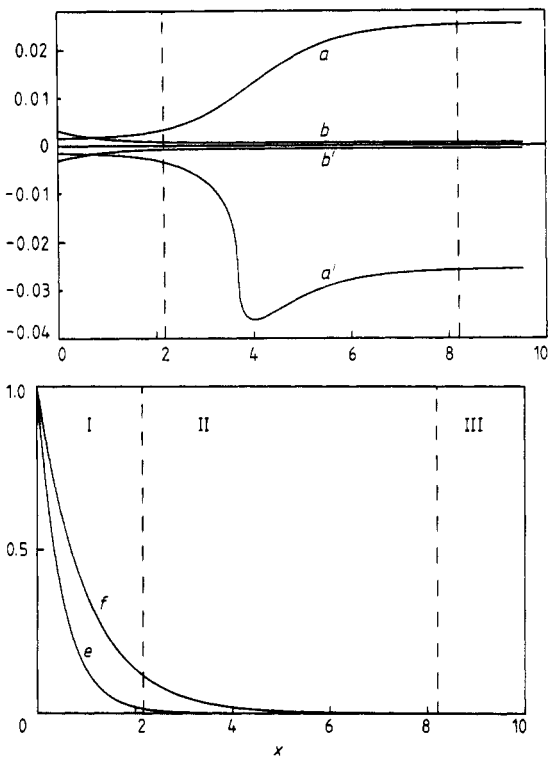

Figure 1. Statistical noise level in a measured autocorrelation function. Curves a, a' and b, b' are standard deviations for noise level in normalised AF(1) and AF(2), respectively, calculated acording to (21). Curves e and f are normalised second- (value 1 subtracted) and first-order autocorrelation functions, respectively. $x = \tau/T_c$, $T_c/T_s = 68$, $\bar{n} = 0.5$, $N = 10^8$.

errors in AF(1) and AF(2). We must use here a convention that $\sqrt{F(\tau)} = -\sqrt{-F(\tau)}$ if $F(\tau) < 0$. From figure 1 we see that there are three characteristic regions and in intensity correlation measurements we should always avoid the second range, for which standard deviation functions $\Delta g^{(1)}_{\pm\sigma}(\tau)$ are not symmetrical, if the transformation from $g^{(2)}$ to $g^{(1)}$ is required.

By setting the proper value for the sample time in a correlator, it is possible to fulfil the first region conditions

$$1 \geqslant (g^{(2)}(\tau) - 1) \geqslant 0.01 \qquad \text{or} \qquad 1 \geqslant g^{(1)}(\tau) \geqslant 0.1 \tag{23}$$

valid when $N \approx 10^8$ and $\bar{n} \approx 0.5$ (in general, the position of the second region and low limits in (23) depend on the number of samples and mean number of counts). This range corresponds to that proposed by Pusey and Vaughan [22], for which we have the maximum information about the decay constant of the AF(1).

When conditions (23) are fulfilled then we can estimate (6) in the first range for a single measurement (relations (12) and (17) are used; $\sqrt{\bar{B}/\hat{B}}$ is replaced by 1)

$$\begin{aligned} \hat{g}^{(1)}_0(\tau) &\approx g^{(1)}(\tau) + \Delta g^{(1)}_r(\tau) \\ &= g^{(1)}(\tau) + \varepsilon_r(\tau)/(2g^{(1)}(\tau)). \end{aligned} \tag{24}$$

5. Power distribution of the noise in the second-order autocorrelation function

5.1. Noise in the intensity of scattered light

AF(2) is formed from intensity fluctuations $n(m)$ in the light scattered on particles moving randomly in the solution (1). Measured intensity fluctuations are influenced by two random processes (two types of noise): photodetection (shot noise) and diffusion (Brownian motion of particles in a measured sample—diffusional noise). In dynamic light scattering experiments, when a single-photon counting method is used, the intensity correlation function and the power density spectrum, PDS, may be expressed as [25-28]:

$$\begin{aligned} g^{(2)}(\tau) &= (1/\bar{n})\delta(\tau) + 1 + |g^{(1)}(\tau)|^2 \\ &= (1/\bar{n})\delta(\tau) + 1 + \int_0^\infty [G(\Gamma') * G(\Gamma')] \exp(-\Gamma'|\tau|)\, d\Gamma' \\ &= (1/\bar{n}) g_{ph}^{(2)}(\tau) + g_d^{(2)}(\tau) \end{aligned} \tag{25}$$

and

$$\begin{aligned} S_i(\omega) &= (1/2\pi)(1/\bar{n}) + \delta(\omega) + (1/\pi) \int_0^\infty \frac{[G(\Gamma') * G(\Gamma')]\Gamma'}{\Gamma'^2 + \omega^2}\, d\Gamma' \\ &= (1/\bar{n}) S_{ph}(\omega) + \delta(\omega) + S_d(\omega). \end{aligned} \tag{26}$$

Here $G(\Gamma')$ is the polydispersity of the measured sample (see [28] and also appendix A) and $\Gamma' = D(r)q^2$ (4); '$*$' denotes a convolution. In the monodisperse case, when $G(\Gamma') = \delta(\Gamma' - \Gamma)$, we have

$$g_d^{(2)}(\tau) = 1 + \exp(-2\Gamma|\tau|) \qquad S_d(\omega) = (2/\pi) \frac{\Gamma}{4\Gamma^2 + \omega^2}. \tag{27}$$

The courses of correlation functions (25) and power density spectra (26) in the monodisperse case (27) are well known in literature (see figures 3 and 4 in [27]). Photodetection noise has constant PDS ('white' noise which corresponds to the correlation function concentrated in one point), but diffusional noise has power density concentrated at low frequencies. PDS can be also obtained from an experiment in which intensity fluctuations are treated by a spectrum analyser in which a bandpass frequency filter has constant $\Delta\omega$ [24]. On the other hand, by keeping constant the quality of the measuring circuit, $Q = \omega/\Delta\omega = \text{constant}$, it is possible to obtain a power spectrum

$$S_i(\omega)\Delta\omega = (P_i(\omega)/\omega)\Delta\omega = P_i(\omega)(\Delta\omega/\omega) = P_i(\omega)(1/Q) \tag{28}$$

which describes a power distribution of different frequency components in light fluctuations (measured signal in spectrum analysis is proportional to $S_i(\omega)\Delta\omega$). From (26), (27) and (28), and using $\omega = 2\pi/T$, we have (by neglecting DC photocurrent)

$$\begin{aligned} \boldsymbol{P}_i(T) &= (1/\bar{n}) \frac{1}{T} + \frac{TT_c}{T^2 + (\pi T_c)^2} \\ &= (1/\bar{n}) \boldsymbol{P}_{ph}(T) + \boldsymbol{P}_d(T) \qquad T_c = 1/\Gamma. \end{aligned} \tag{29}$$

This is a power spectrum, PS, in a period domain and $\boldsymbol{P}_{ph}(T)$, $\boldsymbol{P}_d(T)$ are power spectra for photodetection and diffusional noise respectively. These functions are plotted in

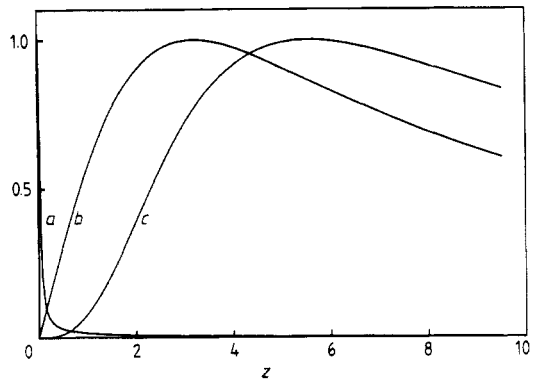

Figure 2. Normalised power spectra (to the maximum = 1) for noise in measured intensity fluctuations in the period domain. Curve *a* is the power spectrum for photodetection noise and curve *b* is the power spectrum for diffusional noise. Curve *c* represents the power spectrum for diffusional noise which is present only in a measured autocorrelation function. Lorentzian case; $z = T/T_c$. Curve *b* has the maximum at $z = \pi$ and curve *c* at $z = \pi\sqrt{3}$.

figure 2 separately, normalised to the maximum equal to 1. Photodetection noise power is mainly concentrated at high frequencies (typical period $2T_s$, T_s = sample time) while diffusional noise has a power maximum at $T_m = \pi T_c$ and its distribution is rather broad.

5.2. Noise in the second-order autocorrelation function

A measured and normalised AF(2) always contains a statistical error (noise) the actual course of which is described by reduced noise fluctuations $\varepsilon_r(\tau)$ (sections 2 and 3). Similarly as for intensity fluctuations, section 5.1, we can obtain a PS for fluctuations $\varepsilon_r(\tau)$ by applying the Fourier transform to the general expression $\Lambda_{\iota\kappa,r}$ from (8) which describes the correlations in $\varepsilon_r(\tau)$. These correlations may be expressed in the form (appendix A):

$$\Lambda_{\iota\kappa,r} = N^{-1}\Big\{(T_c/T_s)\Big[\int_0^\infty \boldsymbol{G}(\Gamma,\tau_\kappa)\exp(-\Gamma(\tau_\iota-\tau_\kappa))\,d\Gamma + 4\int_0^\infty\int_0^\infty \boldsymbol{G}_u(\Gamma,\Gamma',\tau_\kappa)\exp(-(\Gamma+\Gamma')(\tau_\iota-\tau_\kappa))\,d\Gamma\,d\Gamma'\Big] + (1+|g^{(1)}(\tau_\kappa)|^2)\delta(\tau_\iota-\tau_\kappa)/(\bar{n}^2)\Big\} \qquad \tau_\iota \geqslant \tau_\kappa > 0 \tag{30}$$

where $T_c = W^{-1}$ is the effective coherence time [13] and τ_ι and τ_κ are positions of points of in AF(2) for which the correlation $\Lambda_{\iota\kappa,r}$ is evaluated (see figure 3). Distributions $\boldsymbol{G}(\Gamma,\tau_\kappa)$, $\boldsymbol{G}_u(\Gamma,\Gamma',\tau_\kappa)$ and function $|g^{(1)}(\tau_\kappa)|^2$ in (30) depend on the position τ_κ in AF(2) but tend to $\boldsymbol{G}(\Gamma,\infty)$, 0 and 0, respectively, when $\tau_\kappa \to \infty$ (see appendix A). It means then, that a random process $\varepsilon_r(\tau)$ becomes stationary when $\tau_\kappa \gg T_c$ (the tail of AF(2)). From (30), by employing the Fourier transform, we have the PDS for the stationary range of $\varepsilon_r(\tau)$

$$S_\varepsilon(\omega) = N^{-1}\Big[(T_c/T_s)(1/\pi)\int_0^\infty \frac{\boldsymbol{G}(\Gamma,\infty)\Gamma}{\Gamma^2+\omega^2}\,d\Gamma + (1/2\pi)/(\bar{n}^2)\Big]. \tag{31}$$

In the monodisperse case, one can simply compare power spectra for $\varepsilon_r(\tau)$ with those for intensity fluctuations (see section 5.1). Based on (20), we have correlations ($\tau_\iota \geqslant \tau_\kappa \gg 1/\Gamma$)

$$\Lambda_{\iota\kappa,r} = N^{-1}\{(T_c/T_s)(\tfrac{1}{2}+\Gamma(\tau_\iota-\tau_\kappa))\exp(-2\Gamma(\tau_\iota-\tau_\kappa)) + (2/\bar{n})\exp(-2\Gamma(\tau_\iota-\tau_\kappa)) + \delta(\tau_\iota-\tau_\kappa)/(\bar{n}^2)\} \tag{32}$$

and the corresponding PDS

$$S_\varepsilon(\omega) = N^{-1}\left\{(T_c/T_s)(2/\pi)\frac{4\Gamma^3}{(4\Gamma^2+\omega^2)^2} + (2/\pi)(2/\bar{n})\frac{\Gamma}{4\Gamma^2+\omega^2} + (1/2\pi)/(\bar{n}^2)\right\}. \tag{33}$$

For the diffusional noise only ($\bar{n}$ large) we have a PDS

$$S_\varepsilon(\omega) = N^{-1}(2/\pi)(T_c/T_s)\frac{4\Gamma^3}{(4\Gamma^2+\omega^2)^2} \tag{34}$$

or a PS

$$\boldsymbol{P}_\varepsilon(T) = N^{-1}(T_c/T_s)\frac{T_c T^3}{(T^2+(\pi T_c)^2)^2}. \tag{35}$$

$\boldsymbol{P}_\varepsilon(T)$ is plotted in figure 2 (normalised to the maximum = 1) in order to compare it with $\boldsymbol{P}_d(T)$. It has a maximum at $T_m = \sqrt{3}\pi T_c$. In the case when $\bar{n}$ is not very large, the final PS, as we can see from (33) (after transferring this expression to the period domain), is a simple combination of $\boldsymbol{P}_\varepsilon(T)$, $\boldsymbol{P}_d(T)$ and $\boldsymbol{P}_{ph}(T)$.

As was mentioned above, the random process $\varepsilon_r(\tau)$ is not a stationary one, but this non-stationary range is limited only to τ not much larger than T_c, which is also the best for the measurement of AF(2) (see section 4). The determination of PS in this region is a difficult task but it seems that two kinds of approximation are considered. First, we can say that when $\tau_\iota \approx \tau_\kappa$ then $\Lambda_{\iota\kappa,r}$ describes a correlation in a quasistationary process in the non-stationary range of $\varepsilon_r(\tau)$ (random processes in points $\tau_\iota \approx \tau_\kappa$ have nearly the same distributions $\boldsymbol{G}(\Gamma, \tau_\kappa)$, $\boldsymbol{G}_u(\Gamma, \Gamma', \tau_\kappa)$ and nearly the same variances, (21)) and put τ_κ = constant in both distributions and in function $|g^{(1)}(\tau_\kappa)|^2$, occurring in (30), during PS determination. This procedure gives different spectra for each point in the non-stationary range of $\varepsilon_r(\tau)$.

The second approximation, which could be useful in simulations of diffusional noise in AF(2) [31], is based on the assumption that spectra in non-stationary and stationary ranges do not differ too much. This can be easily checked in the monodisperse case by using the first approximation (this is also seen in figure 3, see section 6). The final diffusional spectrum is always a combination of $\boldsymbol{P}_\varepsilon(T)$ and $\boldsymbol{P}_d(T)$ but with different coefficients. The main difference between these two ranges is then in the variance of diffusional fluctuations (21). The second approximation then corresponds to the extension of the stationary spectrum (33) on the whole τ range with the appropriate deviation scale (21).

From figure 2 we can conclude that in real, measured AF(2) we should expect 'fast' fluctuations (fast, compared with the delay timescale (23)) being mainly the result of the photodetection process and rather 'slow' diffusional components with the most probable period T_m lying between πT_c and $\sqrt{3}\pi T_c$. In other words, fluctuations in measured AF(2) look similar to those in the intensity of scattered light (see figure 5 in

an article by Ford in [1], but of course the deviation range for them is much smaller because of the factor N^{-1} in (33)-(35).

6. Experimental

In order to see statistical error courses in AF(2), two series of dynamic light scattering experiments have been performed for the same scattering sample containing a standard Dow latex particle suspension ($\bar{r} = 42.5$ nm). AF(2) were formed in a correlator with 136 channels in full delay time scale (model BI 2020, Brookhaven Instruments). During experiments, the temperature of the sample and the laser power (He–Ne laser, Spectra-Physics, model 124B) were continuously controlled to avoid the long-time instabilities. In order to remove dust particles from the scattering sample, the particle suspension has been filtered with the aid of a 0.2 μm Nuclepore filter. Each of these two series of measurements includes a number of short-time measurements with $N = 2 \times 10^4$ and one long-time measurement with $N = 2 \times 10^8$, but these two series have been performed with different delay timescales; $10T_c$ and $2T_c$ (23) and with different mean number of counts per sample time; $\bar{n} \approx 0.7$ and $\bar{n} \approx 0.5$, respectively. Standard deviations for relative statistical errors in long-time measurements are only 1% of those relative deviations in short-time measurements. For this reason these long-time data, after division of all channel values and baselines by factor 10^4, are very good approximations of expected values for short-time measurements. In figure 3 are shown typical differences $\hat{g}_0^{(2)}(\tau) - g^{(2)}(\tau) = \varepsilon_r(\tau) + \delta B_\varepsilon$ (1), (2), (12) for the first series of measurements (delay timescale $10T_c$), where all AF(2) have been normalised by using the 'expected' baseline taken from long-time measurement. Constant shifts in all correlator channels, δB_ε, and slow frequency components in $\varepsilon_r(\tau)$ are easily seen. In figure 4 are shown differences $\hat{g}^{(2)}(\tau) - g^{(2)}(\tau) \approx \varepsilon_r(\tau) - \delta B \chi^2(\tau)$ (13), (17), realised for the same AF(2) as shown in figure 3, but $\hat{g}^{(2)}(\tau)$ have been normalised by using proper baselines obtained in each short-time experiment. Constant shifts, δB_ε, are fully compensated. Figure 5 illustrates the similar differences as shown in figure 3 but the delay timescale is limited to the non-stationary range of $\varepsilon_r(\tau)$ ($2T_c$ = second series of measurements) and constant shifts $\delta B = \delta B_\varepsilon$ are subtracted. Because of the better separation between photodetection noise distribution $\boldsymbol{P}_{ph}(T)$ and diffusional noise $\boldsymbol{P}_d(T)$ and $\boldsymbol{P}_\varepsilon(T)$ (see figure 2), slow frequency components in $\varepsilon_r(\tau)$ are more easily seen than in figure 3.

7. Statistical errors in AF(2) and a size distribution

Equations (4), (5), (6) and (24) show the relationship between the size distribution $n(r)$ and statistical errors in AF(2). $\Delta g_\varepsilon^{(1)}(\tau)$ has a complicated character, governed by $\varepsilon_r(\tau)$, in which 'slow' components dominate (section 5.2, figures 3 and 5).

Independently of the calculating method used for solving a Fredholm integral equation (4), to determine SD from measured AF(2) by using the Siegert relation (3), 'fast' noise components should be averaged during the calculation procedure but 'slow' deviations should be specially treated by each calculation method because slow changes in AF(2) are responsible for changes in SD. This is explained in figures 6 and 7.

Figure 6 illustrates three size distribution functions (Schulz distribution, [30]) having the following, relative 'widths' Δ (standard deviation/mean value); $\Delta = 10^{-1}$, $\Delta_+ = 1.1 \times 10^{-1}$, $\Delta_- = 0.9 \times 10^{-1}$, with $\bar{r} = 50$ nm.

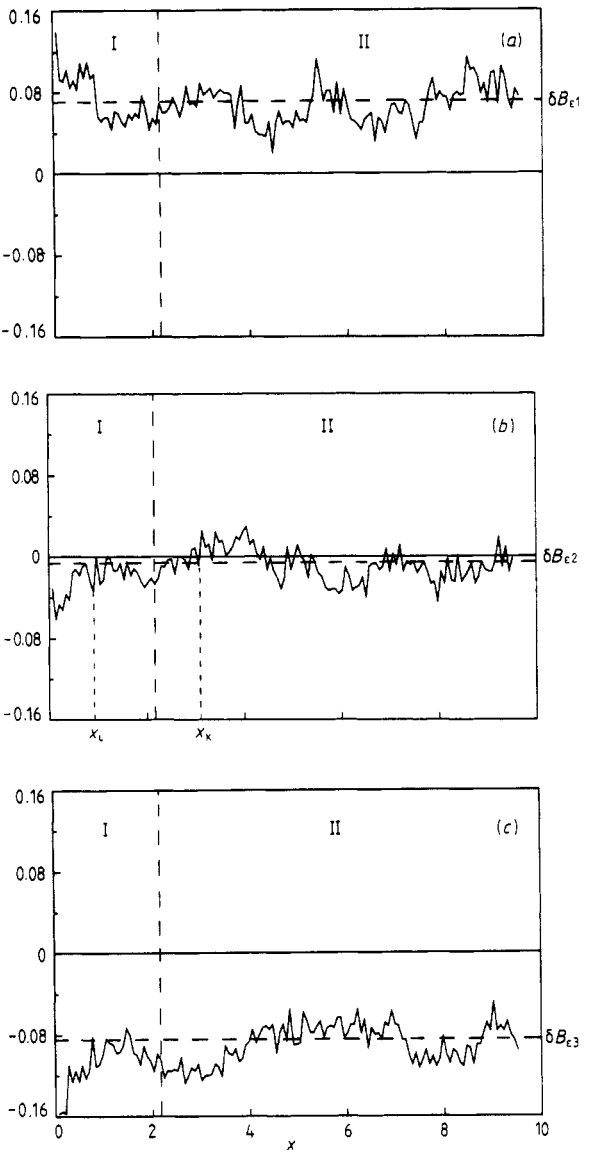

Figure 3. Examples of experimental, statistical error courses in unbiased AF(2) estimators (normalised with $\bar{B}$, see sections 2, 3, 6). Differences between unbiased AF(2) estimators and the expected value, realised for the same scattering sample. I, nonstationary range; II, stationary range.

In figure 7 are shown differences of the corresponding AF(2); $\Delta g_+^{(2)}(\tau) = g_+^{(2)}(\tau) - g^{(2)}(\tau)$, $\Delta g_-^{(2)}(\tau) = g_-^{(2)}(\tau) - g^{(2)}(\tau)$ (normalised), calculated for hollow spheres such as lipid vesicles with the aid of (3) and (4) (program SIMVE, H Ruf, unpublished), by using conditions given in (23).

In addition, this picture describes the differences for AF(2) calculated for SD having the following relative widths: $\Delta = 3.16 \times 10^{-2}$, $\Delta_+ = 3.48 \times 10^{-2}$, $\Delta_- = 2.84 \times 10^{-2}$ and $\Delta = 10^{-2}$, $\Delta_+ = 1.1 \times 10^{-2}$, $\Delta_- = 0.9 \times 10^{-2}$.

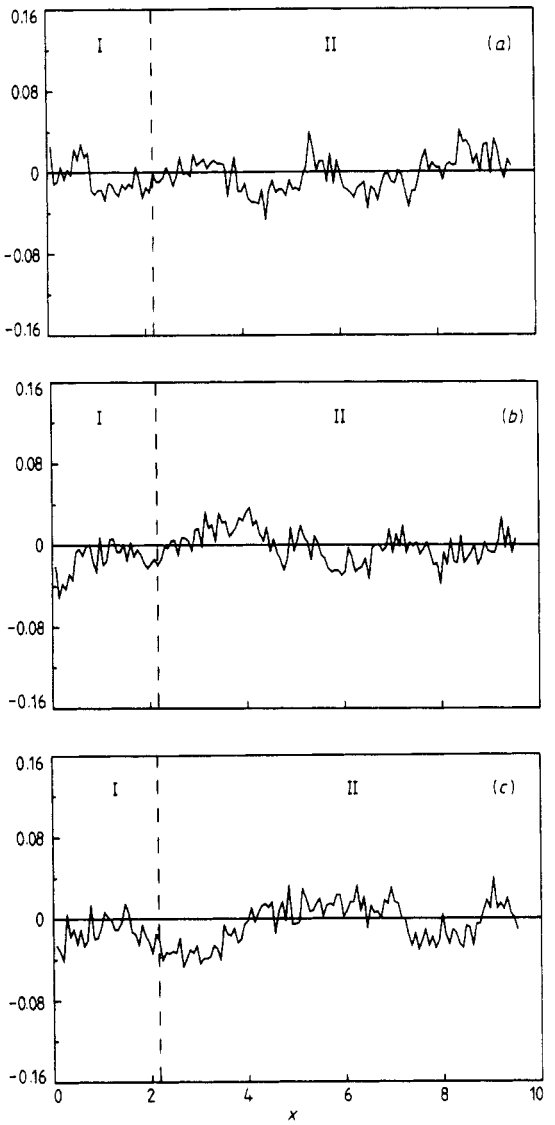

Figure 4. Differences between biased AF(2) estimators (normalised with $\hat{B}$, see sections 2, 3, 6) and the expected value, realised for the same data as given in figure 3. I, nonstationary range; II, stationary range. Calculated baseline errors, $\delta B = (\hat{B} - \bar{B})/\bar{B}$, are 0.0715, −0.0054, −0.0826, respectively.

These three groups of simulated size distributions have the same range of relative width deviation, ±10%, but the corresponding relative deviations in AF(2) are much less and vary from about $\pm 2 \times 10^{-3}$ for the first group of SD to about $\pm 2 \times 10^{-5}$ for the third group of SD representing much narrower size distributions (figure 7).

Correlation functions $g_{+}^{(2)}(\tau)$ and $g_{-}^{(2)}(\tau)$ may be expressed as

$$g_{\pm}^{(2)}(\tau) = g^{(2)}(\tau) + \Delta g_{\pm}^{(2)}(\tau) \tag{36}$$

where $\Delta g_{\pm}^{(2)}(\tau)$ represents deviations shown in figure 7, which relate to an uncertainty

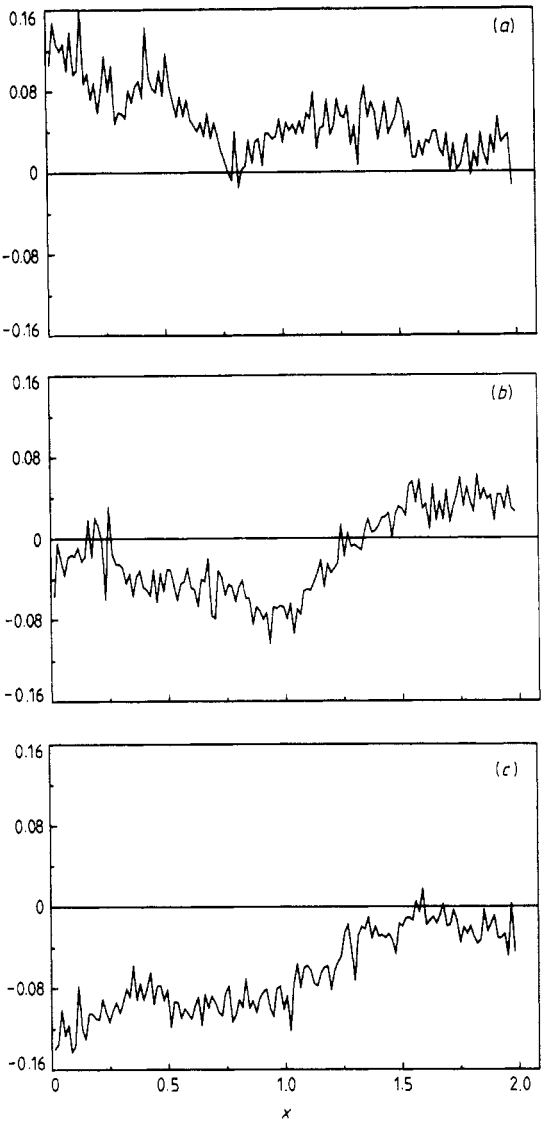

Figure 5. Differences between unbiased AF(2) estimators (normalised with $\bar{B}$; see sections 2, 3, 6) and the expected value, realised for the second series of measurements (delay timescale $2T_c$). Constant shifts $\delta B = \delta B_r$ subtracted.

in the widths of SD. Similarly to (24), by using (3), we can write that

$$g_{\pm}^{(1)}(\tau) \approx g^{(1)}(\tau) + \Delta g_{\pm}^{(2)}(\tau)/(2g^{(1)}(\tau)). \tag{37}$$

From the comparison of the expressions (24) and (37) it would result that the measurement time in a dynamic light scattering experiment should be long enough in order to obtain the standard deviation for statistical errors in AF(2) comparable to (not greater than) the expected uncertainty in AF(2) coming from the expected error in SD width, i.e.

$$\max_{\tau}[\sigma(\tau)] \approx \max_{\tau}|\Delta g_{\pm}^{(2)}(\tau)|. \tag{38}$$

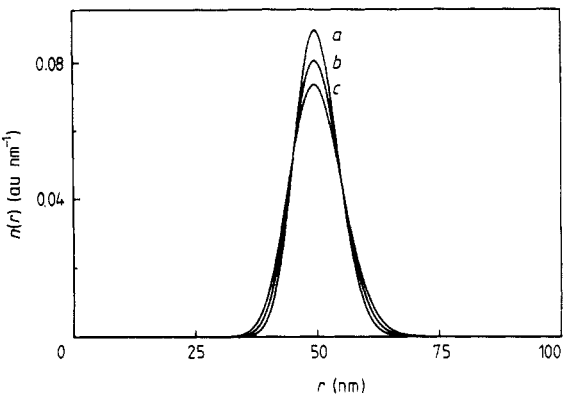

Figure 6. Examples of size distributions of Schulz type, having the following relative widths: *a*: $\Delta_- = 0.9 \times 10^{-1}$, *b*: $\Delta = 10^{-1}$, *c*: $\Delta_+ = 1.1 \times 10^{-1}$.

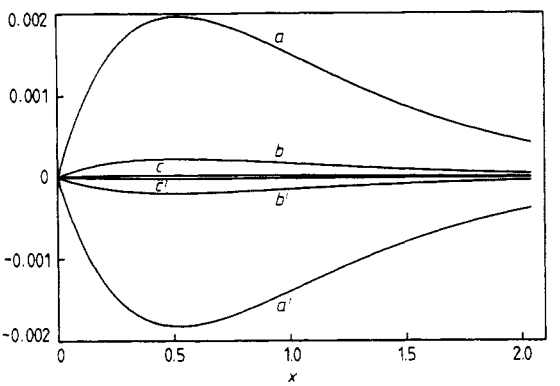

Figure 7. Curves *a*, *a'* represent differences between normalised autocorrelation functions calculated for distributions given in figure 6. Curves *b*, *b'* and *c*, *c'* are the corresponding differences between normalised AF(2) calculated for narrower Schulz distributions (see section 7).

From (38) one can estimate the necessary number of samples N or the necessary measurement time NT_s which should guarantee the expected error in width Δ of SD. In order to determine $[\sigma(\tau)]_{max}$ we can use the variance (21), taking into account the fact that the limits for variance $\Lambda_{\kappa\kappa,r}(\tau_\kappa \to 0, \tau_\kappa \to \infty)$ are nearly the same in mono- and poly-disperse cases in most of experiments (appendix A).

In table 1 are listed the results of such estimations made for four Schulz distributions having various widths Δ but the same width deviation, $\pm 10\%$, and the same $\bar{r} = 50$ nm, under the assumption that $\bar{n} = 0.5$ and $T_c/T_s = 68$.

These results are in a good agreement with the experiment and suggestions of several authors (see the article of Wiener and Tscharnuter in [5]). From table 1 we see that there is a very steep dependence of the necessary number of samples (necessary measurement time) N on Δ, when the same SD width accuracy is expected,

$$N \sim \Delta^{-4} \tag{39}$$

and that very narrow distributions $n(r)$ may in practice be unmeasurable, especially for large $\bar{r}$ values when a large sample time should be used in order to fulfil the

Table 1.

Δ and expected range of deviation of Δ ($\bar{r} = 50$ nm)	$\Delta g_{\pm}^{(2)}(\tau)$ (maximum over τ)	Necessary number of samples N	Necessary measurement time NT_s ($T_s = 10\ \mu$s)
$3.16 \times 10^{-1} \pm 10\%$	$\pm 2 \times 10^{-2}$	$> 6 \times 10^{5}$	>6 s
$10^{-1} \pm 10\%$	$\pm 2 \times 10^{-3}$	$> 6 \times 10^{7}$	>11 min
$3.16 \times 10^{-2} \pm 10\%$	$\pm 2 \times 10^{-4}$	$> 6 \times 10^{9}$	>18 h
$10^{-2} \pm 10\%$	$\pm 2 \times 10^{-5}$	$> 6 \times 10^{11}$	>75 days

Table 2.

Δ and expected range of deviation of Δ ($\bar{r} = 50$ nm)	Level of statistical deviations	Number of samples N	Measurement time NT_s ($T_s = 10\ \mu$s)
$3.16 \times 10^{-1} \pm 0.3\%$	$\pm 6.32 \times 10^{-4}$	$\approx 6 \times 10^{8}$	≈2 h
$10^{-1} \pm 3\%$	$\pm 6.32 \times 10^{-4}$	$\approx 6 \times 10^{8}$	≈2 h
$5.62 \times 10^{-2} \pm 10\%$	$\pm 6.32 \times 10^{-4}$	$\approx 6 \times 10^{8}$	≈2 h
$3.16 \times 10^{-2} \pm 30\%$	$\pm 6.32 \times 10^{-4}$	$\approx 6 \times 10^{8}$	≈2 h
$1.78 \times 10^{-2} \pm 100\%$	$\pm 6.32 \times 10^{-4}$	$\approx 6 \times 10^{8}$	≈2 h
$10^{-2} + 300\%$	$\pm 6.32 \times 10^{-4}$	$\approx 6 \times 10^{8}$	≈2 h
-100%	$\pm 6.32 \times 10^{-4}$	$\approx 6 \times 10^{8}$	≈2 h

conditions given in (23). We may next ask, what is the low limit in size distribution width determination when the measurement time is limited in dynamic light scattering experiments? It seems that a few hours can be the limit of measurement time when avoidance of long-time instabilities in a sample or in a measuring device is desirable. Table 2 lists the expected ranges of deviations in Δ for different SD when the same measurement time, 2 h, is used.

We can thus say, based on the estimation shown in table 2, that for $\bar{r} = 50$ nm ($T_s = 10\ \mu$s) the low limit in SD width determination is about 5×10^{-2}.

8. Conclusions

The Fourier analysis of the correlated noise [13] existing in AF(2) estimators, which is given in section 5, indicates that the statistical error course $\varepsilon_r(\tau)$ can play the role of systematic distortion in one data set (figure 5). Based on the results from section 5, one can estimate that in the monodisperse case half of the most probable period in diffusional fluctuations lies between

$$0.8\,T_{sc} \leqslant 0.5\,T_m \leqslant 1.4\,T_{sc} \tag{40}$$

where $T_{sc} \approx 2\,T_c$ is the optimal delay timescale (23) in a correlator. Unfortunately then, this timescale is the most sensitive for 'slow' components in the diffusional noise.

From table 1 we see that the measurement time, necessary in order to obtain a good result in SD, may be very long for narrow distributions. Thus, one cannot always reach a good result in a few seconds by increasing only the intensity of scattered light.

In this place a strong differentiation between uncorrelated photodetection noise, depending on the intensity of scattered light, and correlated diffusional noise in AF(2), influenced by Brownian motion of particles in a solution and depending on the number of samples N, must be performed. In order to decrease the diffusional noise in AF(2) the number of samples N or the measurement time T_t must be increased (table 1).

It seems, however, that stronger limitations than those given in table 1 should be introduced after taking into account the following facts.

(i) Photodetection noise, the level of which depends on $\bar{n}$, contains mainly 'fast', uncorrelated frequency components (section 5) which should be averaged during the calculation of $n(r)$. This type of noise, however, contributes indirectly to the width of $n(r)$ through the 'smoothing' procedures which are in fact present in each evaluation method (such as the regularisation parameter α in the fitting method, employed in the CONTIN program [7, 12]).

(ii) The shape of 'slow' diffusional components (figure 5) may be very different from deviations $\Delta g_{\pm}^{(2)}(\tau)$ ((36), figure 7) and this may be a source of distortions of a different type such as the appearance of new peaks in $n(r)$ or changes in $\bar{r}$.

(iii) As can be easily checked, relative deviations in AF(2), resulting from deviation in $\bar{r}$, are similar to $\Delta g_{\pm}^{(2)}(\tau)$, which relate to SD width uncertainty, but have a much larger deviation scale. Because of this similarity we can expect that, depending on the actual value of the course of $\varepsilon_r(\tau)$, the final uncertainty in SD determination may 'transfer' between $\bar{r}$ and Δ.

(iv) Estimations given in table 1 relate only to standard deviations for statistical errors. These errors can be much larger in the individual experimental data.

Having a spectrum (33) for experimental noise, one can try to simulate real data [29] and check the influence of different error courses $\varepsilon_r(\tau)$ on the final result. The existence of slow frequency components in $\varepsilon_r(\tau)$ results mainly in size distribution changes and in the appearance of additional distortions, being sometimes similar to those coming from dust particles in a solution. It seems then, that the scientists can see more dust particles in their solutions than really exist. Slow frequency noise components in AF(2) may be the source of additional peaks in final SD [29]. This is one of reasons that investigation of multimodal distributions, by using the dynamic light scattering method, is rather a difficult task [11].

The formal analysis of correlations existing in statistical noise in AF(2) estimators, given in section 3, allows us to interpret individual terms occurring in expressions calculated by Saleh and Cardoso [13] and to find out the proper variance for statistical errors in AF(2) estimators, when size distribution (or linewidth) is determined (section 3, (21)). This analysis indicates also that the so-called 'problem of the normalisation error' does not exist in the normalised dynamic light scattering data because of the 'self-compensation phenomenon'. This surprising result disagrees with the conviction of many people that statistical error in the measured baseline has a significant effect on the final result [5, 17, 23].

Experimental results (section 6, figures 3–5) are, however, in very good agreement with the theory (sections 2, 3 and 5 and figure 2). They indicate also other properties of error courses, which can be concluded from (8), (11) and (15), (18).

(i) The error course deviation in the non-stationary range of $\varepsilon_r(\tau)$ should have, on average, the same sign as the constant shift δB_ε. This is seen in the first and third diagrams of figure 3.

(ii) The term $\delta B \chi_t^2$, resulting from the normalisation and error course deviation in the non-stationary range of $\varepsilon_r(\tau)$, should have opposite signs. These error course

deviations should be then partially compensated when the difference between $\hat{g}(\tau)$ and $g(\tau)$ is calculated. This is seen in the first and third diagrams of figure 4 (non-stationary ranges) while comparing with the same parts of figure 3.

Acknowledgments

This study was carried out at the Max-Planck-Institute for Biophysics (Frankfurt/M, Federal Republic of Germany) and supported by the Max-Planck fellowship. The author wishes to thank Dr E Grell and Dr H Ruf for enabling him to perform and evaluate the light scattering measurements and Dr E Grell, Dr H Ruf, Dr S Provencher and Mr T Claisse for stimulating discussions. Special appreciation is mentioned for support from CPBP 01.06.2.01 (Poland).

Appendix A

Based on the paper of Saleh and Cardoso [13], and taking into account properties of convolution and Laplace transform, we have for the polydisperse case when $\tau_\iota \geqslant \tau_\kappa > 0$ (functions $|\chi_\iota|$, $|y_\iota|$, z_ι, $u_{\iota\kappa}$, in (A1)-(A4) are symmetrical)

$$|\chi_\iota| = \int_0^\infty G(\Gamma) \exp(-\Gamma\tau_\iota)\, d\Gamma \tag{A1}$$

$$|y_\iota| = 2W \int_0^\infty G(\Gamma)\tilde{G}(\Gamma) \exp(-\Gamma\tau_\iota)\, d\Gamma \tag{A2}$$

$$z_\iota = 2W \int_0^\infty G_*(\Gamma)\tilde{G}_*(\Gamma) \exp(-\Gamma\tau_\iota)\, d\Gamma \tag{A3}$$

$$\begin{aligned} u_{\iota\kappa} &= 4W \int_0^\infty \int_0^\infty G_u(\Gamma, \Gamma', \tau_\kappa) \exp(-(\Gamma+\Gamma')(\tau_\iota - \tau_\kappa))\, d\Gamma\, d\Gamma' \\ &= 4W\Bigg[\int_0^\infty \int_0^\infty G_{u1}(\Gamma, \Gamma', \tau_\kappa) \exp(-(2\Gamma+\Gamma')\tau_\kappa) \\ &\quad \times \exp(-(\Gamma+\Gamma')(\tau_\iota - \tau_\kappa))\, d\Gamma\, d\Gamma' \\ &\quad + \int_0^\infty \int_0^\infty G_{u2}(\Gamma.\Gamma', \tau_\kappa) \exp(-\Gamma'\tau_\kappa) \exp(-(\Gamma+\Gamma')(\tau_\iota - \tau_\kappa))\, d\Gamma\, d\Gamma'\Bigg] \end{aligned} \tag{A4}$$

where

$$G_* = G * G \text{ (convolution)} \tag{A5}$$

$$\tilde{G}(\Gamma) = \int_0^\infty G(\Gamma')[1/(\Gamma+\Gamma') + 1/(\Gamma-\Gamma')]\, d\Gamma' \tag{A6}$$

$$\begin{aligned} &G_{u1}(\Gamma, \Gamma', \tau_\kappa) \\ &\quad = G(\Gamma)G(\Gamma') \int_0^\infty \int_0^\infty G(\Gamma'')G(\Gamma''') \\ &\quad \times \left[\frac{1}{\Gamma+\Gamma'+\Gamma''+\Gamma'''} - \frac{1}{\Gamma+\Gamma'+\Gamma''-\Gamma'''}\right] \exp(-\Gamma''\tau_\kappa)\, d\Gamma''\, d\Gamma''' \end{aligned} \tag{A7}$$

and

$$G_{u2}(\Gamma, \Gamma', \tau_\kappa) = G(\Gamma)G(\Gamma') \int_0^\infty \int_0^\infty G(\Gamma'')G(\Gamma''') \times \left[\frac{1}{\Gamma+\Gamma'-\Gamma''+\Gamma'''} - \frac{1}{\Gamma+\Gamma'-\Gamma''-\Gamma'''} \right] \exp(-\Gamma''\tau_\kappa)\, d\Gamma''\, d\Gamma''' \tag{A8}$$

Next, after separating the delay time dependence in all terms occurring in $\Lambda_{\iota\kappa,r}$ (equation (8); see also [13]) for dependence on τ_κ (position) and $\tau_\iota - \tau_\kappa$ (delay), as for example in terms $z_{\iota-\kappa}$ and $z_{\iota+\kappa}$,

$$z_{\iota-\kappa} = 2W \int_0^\infty G_*(\Gamma)\tilde{G}_*(\Gamma) \exp[-\Gamma(\tau_\iota - \tau_\kappa)]\, d\Gamma = 2W \int_0^\infty \boldsymbol{G}_{z-}(\Gamma) \exp[-\Gamma(\tau_\iota - \tau_\kappa)]\, d\Gamma \tag{A9}$$

$$z_{\iota+\kappa} = 2W \int_0^\infty G_*(\Gamma)\tilde{G}_*(\Gamma) \exp[-\Gamma(\tau_\iota + \tau_\kappa)]\, d\Gamma = 2W \int_0^\infty [G_*(\Gamma)\tilde{G}_*(\Gamma) \exp(-2\Gamma\tau_\kappa)] \exp[-\Gamma(\tau_\iota - \tau_\kappa)]\, d\Gamma = 2W \int_0^\infty \boldsymbol{G}_{z+}(\Gamma, \tau_\kappa) \exp[-\Gamma(\tau_\iota - \tau_\kappa)]\, d\Gamma \tag{A10}$$

and by using again properties of convolution and the Laplace transform (for the calculation of products of χ and y) we obtain distribution $\boldsymbol{G}(\Gamma, \tau_\kappa)$ which forms, together with $\boldsymbol{G}_u(\Gamma, \Gamma', \tau_\kappa)$, (30). Distribution $\boldsymbol{G}_{z-}(\Gamma)$ in $z_{\iota-\kappa}$, (A9), and the corresponding distribution $\boldsymbol{G}_{\chi-}(\Gamma)$ in $\chi^2_{\iota-\kappa}$ do not depend on the position τ_κ (figure 3) but distributions in other quantities occurring in $\Lambda_{\iota\kappa,r}$ (8) depend on this position in the similar way as $\boldsymbol{G}_{z+}(\Gamma, \tau_\kappa)$ (A10) or $\boldsymbol{G}_u$ (A4), i.e. they are vanishing when $\tau_\kappa \to \infty$. For this limit we have $\boldsymbol{G}(\Gamma, \infty) = 2\boldsymbol{G}_{z-}(\Gamma) + (2/\bar{n})(T_s/T_c)\boldsymbol{G}_{\chi-}(\Gamma)$.

In order to determine the limits for the variance $\Lambda_{\kappa\kappa,r}(\tau_\kappa \to 0, \tau_\kappa \to \infty)$ in the polydisperse case, we should point out that terms type χ and y have the same limits either in mono- and polydisperse cases. For terms type z and u we have to estimate the quantity $(\tau_\kappa \to 0, \tau_\iota \to \tau_\kappa)$ [13]

$$W \int_{-\infty}^\infty |x(t)|^4\, dt = \left[\int_0^\infty \int_0^\infty G_*(\Gamma) * G_*(\Gamma) \exp(-\Gamma t)\, d\Gamma\, dt \right] \times \left[\int_0^\infty \int_0^\infty G_*(\Gamma) \exp(-\Gamma t)\, d\Gamma\, dt \right]^{-1} = \left[\int_0^\infty (1/\Gamma) G_*(\Gamma) * G_*(\Gamma)\, d\Gamma \right] \left[\int_0^\infty (1/\Gamma) G_*(\Gamma)\, d\Gamma \right]^{-1} = \overline{(1/\Gamma)}_*/\overline{(1/\Gamma)} = c. \tag{A11}$$

By expanding $\overline{(1/\Gamma)}_*$ and $\overline{(1/\Gamma)}$, in terms of moments of the distribution $G_*(\Gamma)$, we have

$$\overline{(1/\Gamma)}_* = (1/(2\bar{\Gamma}))(1+\tfrac{1}{2}m^{(2)}-\tfrac{1}{8}m^{(3)}+\tfrac{1}{6}(\tfrac{1}{8}m^{(4)}+\tfrac{3}{8}(m^{(2)})^2)-\dots) \tag{A12}$$

$$\overline{(1/\Gamma)} = (1/\bar{\Gamma})(1+m^{(2)}-\tfrac{1}{2}m^{(3)}+\tfrac{1}{6}m^{(4)}-\dots)$$

$$m^{(k)} = \overline{(\Gamma-\bar{\Gamma})^k}/(\bar{\Gamma}^k),\ k=1,2,\dots \tag{A13}$$

and hence we have, if the width of SD, Δ, is not very broad, $c \approx \frac{1}{2}$ as for the monodisperse case when $G(\Gamma) = \delta(\Gamma-\Gamma')$. For example, when $\Delta = \sqrt{m^{(2)}} = 0.1$ then c differs from $\frac{1}{2}$ less than 1%.

Appendix B

There exists the need for a commentary of some statements that have been given by Jakeman *et al* [14] and Saleh and Cardoso [13]. It is rather difficult to agree with the statement that 'Drift of a mean photoelectron count rate during the course of a series of experiments ("... for reasons unconnected with the fundamental statistical properties of the signal ..." [14, p 521]) necessitates normalisation of each complete measurement' [14, p 519]. In fact, there is a need for the determination of the baseline value in order to subtract it from a measured AF(2) for further size distribution (or linewidth) determination (2), (4), (5). This must be done also in the case when no such extra drift of the mean photoelectron count rate is expected. On the other hand, the normalisation, in the sense of the division by a number, cannot change the final result, the accuracy of which is determined by the error course $\varepsilon_r(\tau)$ (see sections 2 and 3).

One also cannot agree entirely with the formulation that the results of Saleh and Cardoso [13] are '... a general expression for the correlation between the readings of various channels of a photon counting digital autocorrelator' [13, p 1907], but the sentence 'In this paper, the overall problem of statistical errors in photon counting spectroscopy is formulated ...' [13, p 1897] is valid. The digital reading procedure in dynamic light scattering measurement is simply a single-photon counting and results in uncorrelated photodetection noise in AF(2) but, on the other hand, this procedure may be a source of certain errors (we can call them 'channel correlations') because each channel content represents a mean value taken from counts in a certain finite time range T_s. These errors should depend on the ratio T_s/T_c but not on N.

The problem of such distortions, appearing in digital techniques, has been investigated by several authors [14, 31, 32]. It seems, however, that the final result (Lorentzian, monodisperse case) [32]:

$$g^{(2)}(\tau) = 1+(\sinh^2\gamma/\gamma^2)\exp(-2\Gamma|\tau|) \qquad \gamma = T_s/T_c \tag{B1}$$

does not agree with the experiment because the pre-exponential factor is greater than 1 and tends to infinity with $T_s \to \infty$. The idea of the temporal integration of the measured intensity is valid (equation (4.1) in [32]) but has not been realised to the very end because finally the autocorrelation function $g^{(2)}$ is integrated, with the prior assumption that the spectrum is Lorentzian (equation (4.5) and (4.7) in [32]). One cannot *a priori* assume that the integrated intensity fluctuations have a Lorentzian spectrum. In fact, the temporal integration of $I(t)$ is a kind of low-pass frequency filter which reduces intensity of high frequencies and their power in $S_d(\omega)$ or $P_d(T)$ (27), (29) and two different power spectra, unfiltered (Lorentzian) and filtered, cannot correspond to the same autocorrelation function as shown in (27) and (B1) (the factor $\sinh^2\gamma/\gamma^2$ is here insignificant). This results from the properties of the Fourier transform. If $F(\gamma, |y|)$

($y = \omega T_c/2\pi$) is the filtering function resulting from the temporal integration of $I(t)$ then the intensity autocorrelation function, distorted by the digital reading procedure, may be evaluated by employing the Fourier transform to the reduces power density spectrum of the diffusional noise (26)

$$g^{(2)}(x) = 1 + \mathcal{F}^{-1}[F^2(\gamma, |y|)S_d(y)] \qquad x = \tau/T_c. \tag{B2}$$

In the monodisperse (Lorentzian) case (25) we obtain

$$g^{(2)}(x) = 1 + f(\gamma, x) * e^{-2|x|} \qquad f(\gamma, x) = \mathcal{F}^{-1}[F^2(\gamma, |y|)]. \tag{B3}$$

where '*' is convolution.

References

[1] Dahneke B E (ed) 1983 *Measurement of Suspended Particles by Quasi-Elastic Light Scattering* (New York: Wiley-Interscience)
[2] Provencher S W, Hendrix J, De Maeyer L and Paulussen N 1978 *J. Chem. Phys.* **69** 4273
[3] Horne D S and Dalgleish D G 1985 *Eur. Biophys. J.* **11** 249
[4] Gulari Es, Gulari Er, Tsunashima Y and Chu B 1979 *J. Chem. Phys.* **70** 3965
[5] Provder T (ed) 1987 *Particle Size Distribution* (Washington, DC: American Chemical Society)
[6] Pecora R and Aragon S R 1974 *Chem. Phys. Lipids* **13** 1
[7] Provencher S W 1982 *Comput. Phys. Commun.* **27** 229
[8] Ostrowski N, Sornete D, Parker P and Pike E R 1981 *Opt. Acta* **28** 1059
[9] Chu B, Gulari Es and Gulari Er 1979 *Phys. Scr.* **19** 476
[10] Stock R S and Ray W H 1985 *J. Polym. Sci.* **23** 1393
[11] Stelzer E, Ruf H and Grell E 1983 *Photon Correlations Techniques in Fluid Mechanics* ed E O Schulz-DuBois (Berlin: Springer)
[12] Provencher S W 1982 *Comput. Phys. Commun.* **27** 213
[13] Saleh B E A and Cardoso M F 1973 *J. Phys. A: Math. Gen.* **6** 1891
[14] Jakeman E, Pike E R and Swain S 1971 *J. Phys. A: Gen. Phys.* **4** 517
[15] Schätzel K 1983 *Opt. Acta* **30** 155
[16] Oppenheim A V and Schafer R W 1975 *Digital Processing* (Englewood Cliffs, NJ: Prentice-Hall)
[17] Oliver C J 1981 *Scattering Techniques Applied to Supramolecular and Nonequilibrium Systems* (New York: Plenum Press)
[18] Schulz-DuBois E O 1983 *Photon Correlation Techniques* ed E O Schulz-DuBois (Berlin: Springer)
[19] Siegert A J F 1943 *MIT Lab. Rep.* **465**
[20] Glauber R J 1963 *Phys. Rev.* **131** 2766
[21] Mandel L and Wolf E 1963 *J. Opt. Soc. Am.* **53** 1315
[22] Pusey P N and Vaughan J M 1975 Light Scattering Intensity Fluctuation Spectroscopy *Dielectric and Related Molecular Processes* vol 2, p 48 ed M Davies (London: Chemical Society)
[23] Oliver C J 1973 *Photon Correlation and Light Beating Spectroscopy* (New York: Plenum)
[24] Degiorgio V and Lastovka J B 1971 *Phys. Rev.* A **4** 2033
[25] Mandel L, Sudarshan E C G and Wolf E 1963 *Proc. Phys. Soc.* **84** 435
[26] Mandel L 1964 *Quantum Electronics III (Proc. 3rd Int. Conf., Paris)* ed P Grivet, N Bloembergen (New York: Columbia University Press)
[27] Chu B 1970 *Ann. Rev. Phys. Chem.* **21** 145
[28] Pecora R 1964 *J. Chem. Phys.* **40** 1604
[29] Kojro Z 1990 Influence of statistical errors on size distributions obtained from dynamic light scattering data. Simulation of experimental noise in intensity autocorrelation function, in preparation
[30] Aragon S R and Pecora R 1976 *J. Chem. Phys.* **64** 2395
[31] Pike E R 1969 *Nuovo Cimento* **1** 277
[32] Jakeman E 1973 *Photon Correlation Light Beating Spectroscopy* (New York: Plenum)
[33] Bott S 1983 *Measurement of Suspended Particles by Quasi-Elastic Light Scattering* ed B E Danneke (New York: Wiley-Interscience)
[34] Chu B 1982 *The Application of Laser Light Scattering to the Study of Biological Motion* (New York: Plenum)
[35] Kojro Z and Peters R 1990 Normalisation and statistical noise level in normalised autocorrelation function, in preparation