ELSEVIER

Ultramicroscopy 77 (1999) 37–48

ultramicroscopy

Model-based two-object resolution from observations having counting statistics

E. Bettens[a,*], D. Van Dyck[a], A.J. den Dekker[a], J. Sijbers[a], A. van den Bos[b]

[a]*Department of Physics, RUCA (University of Antwerp), Groenenborgerlaan 171, 2020 Antwerp, Belgium*
[b]*Department of Applied Physics, Delft University of Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands*

Received 28 May 1998; received in revised form 11 December 1998

Abstract

This paper considers two-object resolution from the viewpoint of model fitting theory. The studied experiment consists in counting events, for example, an electron hitting a detector pixel. It is stated that the precision and the accuracy with which the locations of the objects can be estimated will determine the attainable resolution. Two different approaches are followed. For both, the special case of Gaussian peaks is further investigated. The first approach leads to the maximally attainable precision. It is shown that this precision is determined by a certain factor, which is a function of the distance of the peaks, their widths and the number of counts. This factor will be called the resolution factor. The influence of each of the quantities involved is determined by the way they enter this factor. The second approach defines a probability of resolution, i.e., the probability that the maximum likelihood estimates of the locations will be distinct. It is shown that the resolution factor, which resulted from the first approach, also determines the probability of resolution.

Keywords: Resolution; Cramér–Rao lower bound; Model fitting; Maximum likelihood estimator

1. Introduction

The concept of resolution, the ability to distinguish details, is an important quality measure for imaging systems. Higher resolution often means that the interpretation of the image is easier. In the past, many resolution criteria have been proposed for two-point resolution, i.e., the ability to resolve two adjacent points. These criteria are mostly used in diffraction limited systems: systems where the image of a point is spread by a point-spread function. One of the earliest and most famous criteria is that of Rayleigh [1]. According to Rayleigh, resolution is limited by the width of the main lobe of the point spread function. However, this criterion is only based on the limitations of the human visual system, and does not take into account, for example, the presence of noise. Another classical criterion is that of Rose [2], which approaches resolution in terms of dose, i.e., the number of counts per area. An extensive survey on the concept of resolution can be found in [3]. In the present paper, two-object resolution is studied, i.e., the ability to resolve two objects of equal size and intensity. A global approach is followed: to investigate the

* Corresponding author.

0304-3991/99/$ – see front matter
PII: S0304-3991(99)00006-6

attainable resolution, the size of the objects, the distance between them, as well as the dose are taken into account. It is assumed that a mathematical model for the objects exists and that this model is known, for example, the objects could be atoms, described by identical gaussian probability density functions. The locations of the objects are unknown and appear in the model as parameters. The ultimately attainable resolution is then achieved by using model-fitting techniques: the quality of the estimates of the location parameters will determine the attainable resolution. In this paper we follow two different approaches to investigate this quality.

The first approach is based on the available asymptotic parameter estimation theory. A survey of this theory can be found in handbooks on statistics and parameter estimation, for example, in [4,5]. With the aid of this theory we deduce a lower limit on the attainable precision an unbiased estimator of the distance between the peaks can achieve. It is shown that this precision is determined by a factor that is a function of the distance of the peaks, their widths and the number of counts. This function will be called the resolution factor. The influence of each of the quantities involved is determined by the way they enter this factor. From this factor it can be seen, for instance, that the variance of an unbiased estimator will grow drastically if the distance between the peaks is decreased below a critical value. As a result, the estimated locations will no longer have meaningful values.

In the second approach, we investigate the possible estimates given by the maximum likelihood estimator, one of the most important estimators. It turns out that, for two closely located objects, collapse of the two objects can occur, i.e., for particular sets of observations, the estimated distance between the objects is exactly zero. This is due to a change of structure in the maximum likelihood criterion under the influence of the observations. This kind of structure change has been described in [6–9] for the special case of a least-squares estimator. A rule is derived to calculate the probability that such a collapse will occur (this probability for resolution is similar to the one described in [10] for least-squares estimators).

In these two different approaches, the first related to precision (Standard deviation) and the second to accuracy (bias), the same expression, containing the total number of counts and the distance and width of the objects, appears. The paper is organized as follows. Section 2 gives a short introduction on the theory of parameter estimation and model fitting, which is used in this paper. In Section 3, the attainable precision of an unbiased estimator of the distance parameter is studied. In Section 4, the behavior of the maximum likelihood estimator is studied with the aid of catastrophe theory. In Section 5, a number of numerical examples are discussed.

2. Parameter estimation and model fitting

Consider an experiment that consists of counting events, for example, an electron hitting a pixel. The events are distributed over a number of intervals, described by $\{x_i, i \ldots, M\}$, by a probability density function (pdf). The observations are given by $\{n_i, i = 1, \ldots, M\}$, where n_i describes the number of counts or events in the interval x_i. The total number of counts is defined by N, with $N = \sum_{i=1}^{M} n_i$. The probability that an event occurs in the interval x_i will be denoted by $p(x_i; \theta)$, which is a pdf depending on the parameters of the model $\theta = (\theta_1, \ldots, \theta_K)^t$. This means that the observations are described as

$$\underline{n_i} = Np(x_i; \theta) + \underline{\varepsilon_i} \tag{1}$$

with $\underline{n_i}$ and $\underline{\varepsilon_i}$ stochastic variables; $\underline{\varepsilon_i}$ is the deviation from the expectation value $E[\underline{n_i}]$, which is given by

$$E[\underline{n_i}] = Np(x_i; \theta). \tag{2}$$

These $E[\underline{n_i}]$ form the model of the observations: they describe the expected outcome of the experiment. Model fitting means that the parameters θ must be estimated so that the model fits the observations as well as possible. If there were no noise in the observations, and the model correct, these parameters could be calculated exactly. This is, however, never the case in a real experiment.

There are various ways of estimating the parameters. One of the most important is the maximum likelihood method (ML-method). The likelihood of a set of parameters is defined by the joint probability that the observations occur given this set of parameters. The ML-method consists of maximiz-

ing the likelihood function with respect to the parameters. In an experiment, where we have a total of N counts, the likelihood function L for a set of observations $\{n_i, i = 1, \ldots, M\}$ is given, using multinomial statistics, by

$$L = N! \prod_i \frac{p^{n_i}(x_i; \theta)}{n_i!}. \tag{3}$$

One important property of the ML-method is related to the *Cramér–Rao lower bound* (CRLB) or the *minimum variance bound*, which is a lower bound on the variance of any unbiased estimator [5]. The CRLB does not depend on the method of estimation. If we define $\hat{t} = (\hat{t}_1(\{n_i\}), \ldots, \hat{t}_K(\{n_i\}))^\tau$ as the unbiased estimators of the parameters $\theta = (\theta_1, \ldots, \theta_K)^\tau$, $\mathrm{cov}(\hat{t}, \hat{t})$ as the covariance matrix of $\hat{t}$, with as diagonal elements the variances, and M as the matrix with elements defined as

$$M_{kl} = -E\left[\frac{\partial^2 \ln L}{\partial\theta_k \partial\theta_l}\right] \tag{4}$$

with $k, l = 1, \ldots, K$, the Cramér–Rao inequality states, under certain regularity conditions, that

$$\mathrm{cov}[\hat{t}, \hat{t}] \geqslant M^{-1}. \tag{5}$$

In other words, the difference between the positive left-hand and right-hand members is positive semi-definite. A property of positive semi-definite matrices is that their diagonal elements cannot be negative. Therefore, the diagonal elements of $\mathrm{cov}[\hat{t}, \hat{t}]$, that is, the variances of the estimator $\hat{t}$, must be larger than or equal to the corresponding diagonal elements of M^{-1}. The matrix M^{-1} is the CRLB. The CRLB of functions of the parameters follows relatively easily from the CRLB for the parameters. Let $\hat{r} = (\hat{r}_1(\{n_i\}), \ldots, \hat{r}_L(\{n_i\}))^\tau$ be an unbiased estimator of the functions $(\rho_1(\theta), \ldots, \rho_L(\theta))^\tau$; it can be shown that

$$\mathrm{cov}[\hat{r}, \hat{r}] \geqslant \frac{\partial\rho}{\partial\theta^\tau} M^{-1} \frac{\partial\rho^\tau}{\partial\theta}, \tag{6}$$

where $\partial\rho/\partial\theta^\tau$ is the $L \times K$ Jacobian matrix with as (p, q)th element $\partial\rho_p/\partial\theta_q$ [4]. Under general conditions, ML-estimators attain the CRLB asymptotically, that is, for a large number of observations. Furthermore, it is known that, if an unbiased estimator having a variance equal to the CRLB exists, this estimator is the ML estimator.

In the next sections, we will work with the following discrete pdf, which describes the observations of two identical one-dimensional, real, even objects centered around the location a_1 and a_2:

$$p(x_i; a_1, a_2) = \tfrac{1}{2}(f(x_i - a_1) + f(x_i - a_2)). \tag{7}$$

If an object itself is described by a continuous pdf $g(x)$, the discrete function $f(x_i)$ is given by

$$f(x_i) = \int_{x_i - \Delta/2}^{x_i + \Delta/2} g(x)\,\mathrm{d}x, \tag{8}$$

where Δ is the width of the intervals in which the measurements were made.

3. Resolution in terms of statistical precision

We want to establish how close two objects in an experiment may get, before they can no longer be separated. If there were no noise present, one could resolve the objects even when they were extremely close together. However, in reality there is almost always a certain amount of noise present; consequently, the resolvability of the objects will be limited. The question to be answered then is: when are two objects resolved? One possible approach to this problem is looking at resolution in terms of the precision, that is, the standard deviation, of the estimated parameters.

3.1. Cramér–Rao lower bound for a two-object model

The attainable precision for estimated parameters is determined by Eqs. (4) and (5). The logarithm of L is given by

$$\ln L = \sum_i n_i \ln p(x_i; \theta) + c, \tag{9}$$

where c is a term not depending on the parameters.

The elements of the matrix M, see Eq. (4), are given by

$$M_{kl} = -\sum_i E[\underline{n_i}] \left(\frac{\partial^2 \ln p_i}{\partial a_k \partial a_l}\right)_{a_1^0,\, a_2^0}, \quad k = 1, 2, \quad l = 1, 2 \tag{10}$$

with $p_i = p(x_i; a_1, a_2)$ and a_1^0, a_2^0 the true values of the parameters. It is clear that $M_{12} = M_{21}$, so that

$$\mathrm{Var}_{\mathrm{cr}}(\hat{a}_1) = \frac{M_{22}}{M_{11}M_{22} - M_{12}^2}, \tag{11}$$

$$\mathrm{Var}_{\mathrm{cr}}(\hat{a}_2) = \frac{M_{11}}{M_{11}M_{22} - M_{12}^2}, \tag{12}$$

where $\mathrm{Var}_{\mathrm{cr}}(\hat{a}_i)$ denotes the minimal variance, determined by the CRLB, for any unbiased estimator $\hat{a}_i$ of the parameter a_i. It is the distance between the objects that will determine resolution. Therefore, the location parameters are transformed into a new set of parameters, defined by

$$A_1 = a_1 + a_2, \tag{13}$$

$$A_2 = a_1 - a_2, \tag{14}$$

where A_1, can be regarded as two times the center of mass and A_2 as the distance between the objects. The minimal variances of these new parameters can easily be calculated by using Eq. (6). One finds

$$\mathrm{Var}_{\mathrm{cr}}(\hat{\mathrm{A}}_1) = \frac{M_{11} + M_{22} - 2M_{12}}{M_{11}M_{22} - M_{12}^2}, \tag{15}$$

$$\mathrm{Var}_{\mathrm{cr}}(\hat{\mathrm{A}}_2) = \frac{M_{11} + M_{22} + 2M_{12}}{M_{11}M_{22} - M_{12}^2}, \tag{16}$$

Eq. (16), i.e., the minimal variance of the distance estimator determines the attainable resolution.

What we would like to find, however, is a rule of thumb, that is, an expression that is easy to calculate and to interpret. To accomplish this, a number of approximations are made. Simulations, described in Section 5, justify these approximations.

The first assumption is that M_{11} and M_{22} are almost equal, so that substituting M_{11} for M_{22} will have little influence on the result. This is justified if the objects, which have been assumed to be symmetric, are situated for the most part within the region of observation (it is not justified if, for instance, only one half of an object is measured).

Eqs. (15) and (16) can then be rewritten as

$$\mathrm{Var}_{\mathrm{cr}}(\hat{\mathrm{A}}_1) \approx \frac{2}{M_{11} + M_{12}}, \tag{17}$$

$$\mathrm{Var}_{\mathrm{cr}}(\hat{\mathrm{A}}_2) \approx \frac{2}{M_{11} - M_{12}}. \tag{18}$$

We will now study these variances for objects close together. It then follows from Eqs. (2) and (10) that

$$M_{11} + M_{12} \approx \frac{N}{2} \sum_i \left(\frac{1}{p_i} \left(\frac{\partial p_i}{\partial a_1} + \frac{\partial p_i}{\partial a_2} \right)^2 \right)_{a_1^0,\, a_2^0}, \tag{19}$$

$$M_{11} - M_{12} \approx \frac{N}{2} \sum_i \left(\frac{1}{p_i} \left(\frac{\partial p_i}{\partial a_1} - \frac{\partial p_i}{\partial a_2} \right)^2 \right)_{a_1^0,\, a_2^0}, \tag{20}$$

where, again, M_{11} has been substituted for M_{22} and use has been made of the identity $\sum_i p_i = 1$. We rewrite $\partial p_i/\partial a_1$ and $\partial p_i/\partial a_2$, using Eqs. (7), (13) and (14), as

$$\frac{\partial p_i}{\partial a_1} = -\tfrac{1}{2} f^{(1)}(x_i - (A_1 + A_2)/2), \tag{21}$$

$$\frac{\partial p_i}{\partial a_2} = -\tfrac{1}{2} f^{(1)}(x_i - (A_1 - A_2)/2), \tag{22}$$

where $f^{(1)}(x_i)$ is the first-order derivative of $f(x_i)$ with respect to x_i. Eqs. (21) and (22) can be Taylor expanded about $A_2 = 0$, for peaks close together, as follows:

$$\frac{\partial p_i}{\partial a_1} = -\tfrac{1}{2} \left(f^{(1)}(x_i - a) - \frac{A_2}{2} f^{(2)}(x_i - a) + \cdots \right), \tag{23}$$

$$\frac{\partial p_i}{\partial a_2} = -\tfrac{1}{2} \left(f^{(1)}(x_i - a) + \frac{A_2}{2} f^{(2)}(x_i - a) + \cdots \right) \tag{24}$$

with $a = A_1/2$ and $f^{(2)}(x_i)$ the second-order derivative of $f(x_i)$ with respect to x_i. This leads to

$$\left(\frac{\partial p_i}{\partial a_1} + \frac{\partial p_i}{\partial a_2} \right)^2 \approx (f^{(1)}(x_i - a))^2, \tag{25}$$

$$\left(\frac{\partial p_i}{\partial a_1} - \frac{\partial p_i}{\partial a_2} \right)^2 \approx \frac{A_2^2}{4} (f^{(2)}(x_i - a))^2. \tag{26}$$

Now we find for Eqs. (19) and (20):

$$M_{11} + M_{12} \approx \frac{N}{2} \sum_i \left(\frac{(f^{(1)}(x_i - a))^2}{f(x_i - a)} \right)_{a_1^0,\, a_2^0}, \tag{27}$$

$$M_{11} - M_{12} \approx \frac{N}{8} A_2^2 \sum_i \left(\frac{(f^{(2)}(x_i - a))^2}{f(x_i - a)} \right)_{a_1^0,\, a_2^0}, \tag{28}$$

where $p(x_i; a_1, a_2)$ is approximated by $f(x_i - a)$, since we consider the case where the two objets nearly overlap completely.

3.2. Special case: sum of two Gaussian peaks

Let us now consider the special case where the model is the sum of two Gaussian peaks. Define

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma_0} \exp\left(-\frac{x^2}{2\sigma_0^2} \right). \tag{29}$$

If the intervals of the observations are small enough – in Section 5 we will study the effect of the width of the intervals – $f(x_i)$ is well approximated by $g(x_i)\Delta$ so that the sums in Eqs. (27) and (28) can be approximated by an integral, which gives

$$\sum_i \frac{(g^{(1)}(x_i - a))^2}{g(x_i - a)} \Delta \approx \frac{1}{\sigma_0^2}, \tag{30}$$

$$\sum_i \frac{(g^{(2)}(x_i - a))^2}{g(x_i - a)} \Delta \approx \frac{2}{\sigma_0^4}. \tag{31}$$

Finally, we find for Eqs. (17) and (18)

$$\mathrm{Var}_{\mathrm{cr}}(\hat{A}_1) \approx \frac{4\sigma_0^2}{N}, \tag{32}$$

$$\mathrm{Var}_{\mathrm{cr}}(\hat{A}_2) \approx \frac{8\sigma_0^4}{NA_2^2}. \tag{33}$$

Eq. (33) gives an approximation of the Cramér–Rao variance for the distance parameter, or, in other words, the ultimate precision any unbiased estimator of the distance can achieve. This variance is a function of the total number of counts, the actual distance between the peaks and the width of the gaussian peak used in the model. A higher precision can be achieved by increasing the number of counts, which means increasing the signal-to-noise ratio, as one would expect. Also, if the widths of the peaks are smaller, or if the peaks are further separated, the attainable precision will be better. When the variance of the distance parameter becomes too large, resolution is no longer possible.

The ratio of the standard deviation of the estimated distance to the distance itself, which we define as the resolution factor R, is an important quantity. Using Eq. (33), R is found to be given by

$$R = \frac{\sqrt{\mathrm{Var}_{\mathrm{cr}}(\hat{A}_2)}}{A_2} = 2\sqrt{2}\,\frac{\sigma_0^2}{\sqrt{N}A_2^2}. \tag{34}$$

An obvious resolution criterion is that R must be smaller than 1, that is,

$$\frac{\sigma_0^2}{\sqrt{N}A_2^2} < \frac{1}{2\sqrt{2}} \tag{35}$$

which is a simple and useful rule of thumb. The next question to be answered is whether there exists an unbiased estimator that attains the CRLB. As stated in the previous section, this estimator would be the ML-estimator, which will be studied in the next section.

4. Behavior of the ML-estimator

In this section, the behavior of the ML-estimator is investigated if the distance between the objects becomes small. In simulated experiments, it was noticed that there were two distinct possibilities: the ML-estimates for the locations of two objects coincided, meaning that their estimated distance was zero, or the ML-estimates were distinct, meaning that there were two equivalent maxima of the likelihood function as a function of the distance (equivalent because of the symmetry of the model). This means that the structure of the likelihood function, as a function of the distance, changed from one maximum, located at zero, to two non-zero maxima. This kind of structure change is described by catastrophe theory [11].

4.1. Taylor expansion

Since a complete description of the method is outside the scope of this paper, the method will only be sketched. For a detailed description of the operations involved we refer to [11].

Catastrophe theory is concerned with the structural change of a parametric function under influence of its parameters. It tells us that a structural change of the function is always preceded by a degeneracy of one of its stationary points. The theory also shows that the independent variables of the

function can be split into essential and inessential variables. The essential variables correspond to the directions in which degeneracy may occur. The inessential variables do not play a role at all in the structural change. Further, the structural change of a function is fully described by a structural change of its Taylor polynomial in the essential variable up to and including the first term having a non-vanishing coefficient. Now, suppose that $\hat{a}$ is the ML-solution for the one-component model $f(x_i; a)$, then it can easily be shown that $(\hat{a}, \hat{a})$ is a stationary point of L for the two-component model. This point is called the one-component stationary point. The loglikelihood function $L' = \sum_i n_i \ln p_i$ is then Taylor expanded about this stationary point. The parameters a_1 and a_2 are transformed into the parameters A_1 and A_2, as defined in the previous section. It can then be shown that there is only one essential variable, namely A_2. The coefficient of A_1^2 is always non-zero and negative. Therefore, the variable A_1 is inessential while at stationary points the likelihood function is, in the direction of A_1, always maximal. After a number of calculations, similar to the ones described in [6], the Taylor polynomial in this essential variable describing the structural change, is found to be

$$BA_2^2 + CA_2^4, \tag{36}$$

where the constant term has been omitted and where the coefficients B and C are given by

$$B = \sum_i \frac{n_i}{2} \left(\frac{\partial^2 \ln p_i}{\partial A_2^2} \right)_{\hat{b}}, \tag{37}$$

$$C = \sum_i \frac{n_i}{4!} \left(\frac{\partial^4 \ln p_i}{\partial A_2^4} \right)_{\hat{b}} - \left(\sum_i \frac{n_i}{3} \left(\frac{\partial^3 \ln p_i}{\partial A_1 \partial A_2^2} \right)_{\hat{b}} \right)^2 \Bigg/ \sum_i \frac{n_i}{8} \left(\frac{\partial^2 \ln p_i}{\partial A_1^2} \right)_{\hat{b}} \tag{38}$$

with $\hat{b} = (2\hat{a}, 0)$ the one-component stationary point expressed in terms of the parameters A_1 and A_2. It can be shown that the coefficient C is always negative for commonly used peaklike object functions, e.g., Gaussian functions. Similarly, it can be proven that the coefficient B can be negative as well as positive. Hence, there are two possible cases: if the coefficient $B \leqslant 0$, polynomial (36) has only one stationary point : a maximum at $A_2 = 0$; if $B > 0$, the polynomial has two distinct maxima symmetrically located about zero. Thus if the sign of B goes from a positive value to a negative value, a change of structure in the ML-criterion occurs: it goes from two maxima to one maximum that is always located at zero. If the estimated distance is zero, the two objects collapse, which means that they no longer can be separated. The sign of the coefficient B therefore decides whether the two objects can be separated or not.

4.2. Probability of resolution

The coefficient B depends on the observations n_i, so, if its probability density function is known, the probability of finding $B > 0$, which we define as the probability of resolution, is also known. If the *Central Limit Theorem* applies, which will be assumed, B has a normal distribution, so that

$$P(B > 0) = 1 - P\left(X < \frac{-E[B]}{\sigma_B} \right), \tag{39}$$

where X has a standard normal distribution. The ratio $E[B]/\sigma_B$ plays a crucial role, it will determine the probability that $B > 0$, that is, the probability that the objects will be separated.

4.3. Special case: sum of two Gaussian peaks

We will now consider the pdf of B for Gaussian peaks. First, $E[B]$ and σ_b^2 are calculated. Because $E[n_i] = Np_i$, it follows from Eq. (37) that the expectation value of B is given by

$$E[B] = \frac{N}{2} \sum_i p_i \left(\frac{\partial^2 \ln p_i}{\partial A_2^2} \right)_{\hat{b}}. \tag{40}$$

As before, it will be assumed that the intervals are small enough so that the sum can be replaced by an integral. For Gaussian peaks this leads to

$$E[B] = \frac{NA_2^2}{32\sigma_0^4}, \tag{41}$$

Similarly, the variance is found to be

$$\sigma_B^2 = \frac{N}{32\sigma_0^4} \tag{42}$$

where terms with the factor A_2^2/σ_0^6 are neglected because we consider the case of peaks close together. Then,

$$\frac{E[B]}{\sigma_B} = \frac{\sqrt{N}}{4\sqrt{2}} \frac{A_2^2}{\sigma_0^2}, \tag{43}$$

The larger this ratio, the larger the probability of finding a positive B-value and the larger the probability of separating the peaks. Using Eqs. (43) and (34), Eq. (39) can be written as

$$P(B > 0) = \tfrac{1}{2} + \tfrac{1}{2}\mathrm{erf}\left(\frac{1}{2\sqrt{2}R}\right), \tag{44}$$

where erf() is the error function. Then, the value of $R = 1$, used in Section 3.2, corresponds to $P(B > 0) = 0.69$. Again, as in Section 3, the smaller the factor R, the better the attainable resolution, which means here, the higher the probability of separating the peaks.

For peaks close together, so that the ratio A_2/σ_0 is much smaller than N, a substantial part of the estimates of the distance will be zero. It can be concluded that the ML-estimator is a biased estimator for peaks close together. This means that there does not exist an unbiased estimator that will reach the CRLB. However, this does not imply that there does not exist an unbiased estimator. It is still possible that such an estimator exists, but with a variance large than the CRLB.

5. Discussions and experiments

First, the CRLB of the distance parameter, Eq. (16), and its approximation, Eq. (33), will be investigated for Gaussian peaks, by means of an example. The resolution factor R, i.e., the standard deviation of A_2 divided by A_2 itself (sometimes called the relative error), is the most interesting quantity to investigate. We define R^{exact} as the non-approximated factor R, i.e., the root of Eq. (16), divided by A_2. Fig. 1 shows R^{exact} and R as a function of the ratio A_2/σ_0, Fig. 1(a), and as a function of the number of counts N, Fig. 1(b). From these figures, it is concluded that the approximation predicts the exact value sufficiently accurately to be usable as a rule of thumb, which was the purpose. In Fig. 1(a), where $N = 5000$, it can be seen that the factor $R = 1$ when $A_2/\sigma_0 = 0.2$, that is, when σ_0 is five times large than the distance A_2. It can also be observed that, if, for instance, σ_0 is three times larger than A_2, the factor R is already reduced to approximately 0.4. From Fig. 1(b), where $A_2/\sigma_0 = 0.3$, similar conclusions can be drawn. It can be seen that the influence of A_2/σ_0 is stronger than the influence of N (A_2/σ_0 appears quadratically in Eq. (34), while N appears as a root). However the number of counts is also important, it can make the difference between resolution or not.

Fig. 2 shows the approximation of the CRLB for the distance parameter, as a function of the width of the intervals, but now with the sum in Eq. (31) numerically calculated and not approximated by an integral. It is seen that the CRLB increases only slightly with increasing size of the intervals, with all other quantities kept constant. Increasing the number of intervals, i.e., decreasing the size of the intervals, thus results in a decrease of the CRLB. However, no precision can be gained significantly, by going, for example, from interval size 1 to interval size 0.1. The CRLB reaches the value 36 asymptotically, which is the value given by Eq. (33), where the sum was replaced by an integral. If the size of intervals becomes relatively large, the real CRLB will deviate from the approximation given by Eq. (33), and Eq. (31) should be calculated instead.

In the next experiment 10 000 different sets of observations have been simulated, but all have been generated by the same pdf: sum of two Gaussians at a distance 6 from each other. The width of the Gaussians was 30 and the total number of counts 5000. There are two possibilities: for certain sets of observations a loglikelihood function is found, defined by Eq. (9), that has two distinct maxima as a function of the distance, located symmetrically around zero, depicted in Fig. 3(a). For other sets, the loglikelihood function has only one maximum located at zero, depicted in Fig. 3(b). The dotted line in the figure is the Taylor polynomial.

Eq. (39) predicts that about 69% of the B-values will be positive, meaning that for 31% of the experiments the ML-estimated distance will be zero. We have simulated 10 000 experiments and found a value of 31.5% of negative B-values. The 95% confidence interval contains the predicted

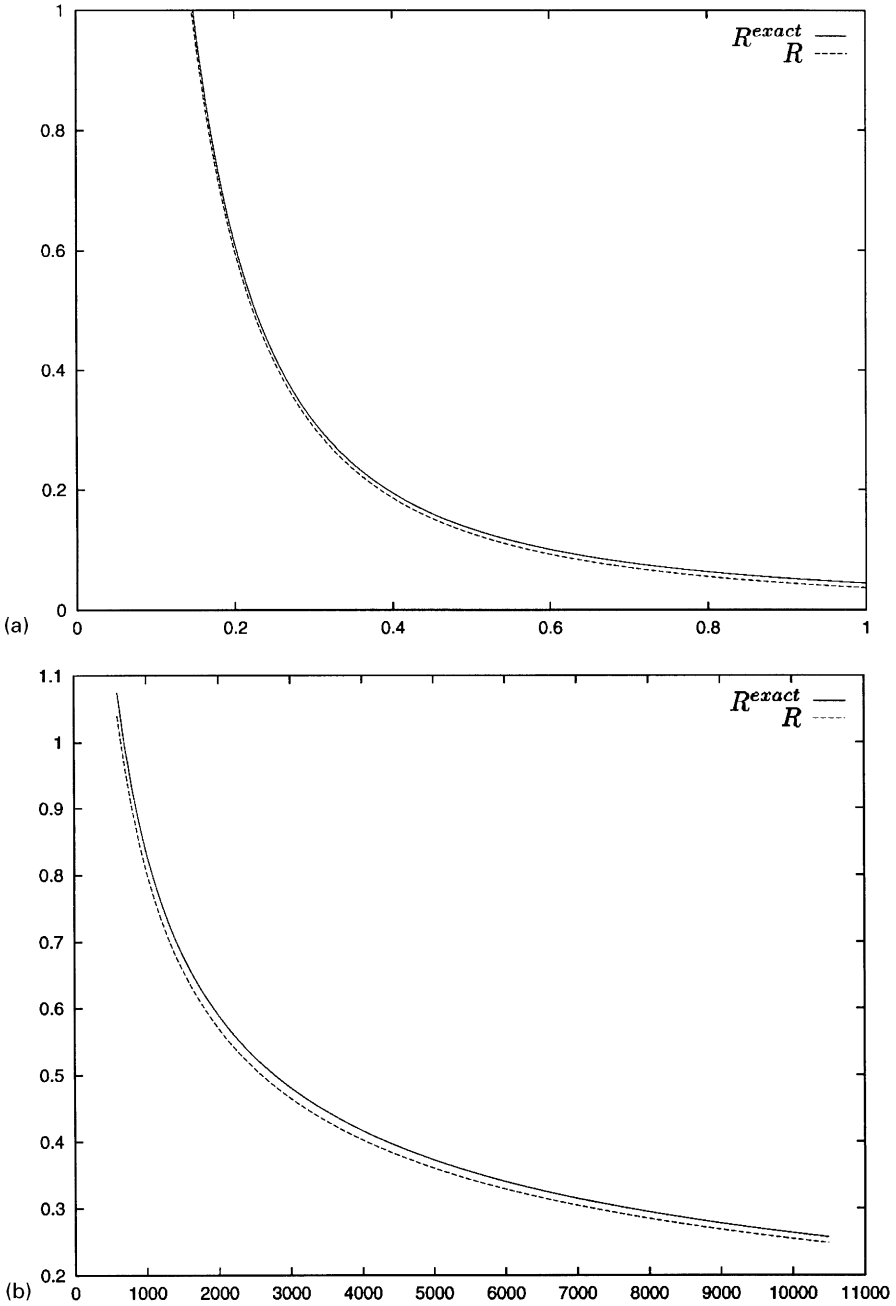

Fig. 1. The resolution factor R^{exact} (exact standard deviation divided by A_2) and its approximation $R(=2\sqrt{2}\sigma_0^2/(NA_2^2))$ as a function of A_2/σ_0, with $N = 5000$ (a), and as a function of N, with $A_2/\sigma_0 = 0.3$ (b).

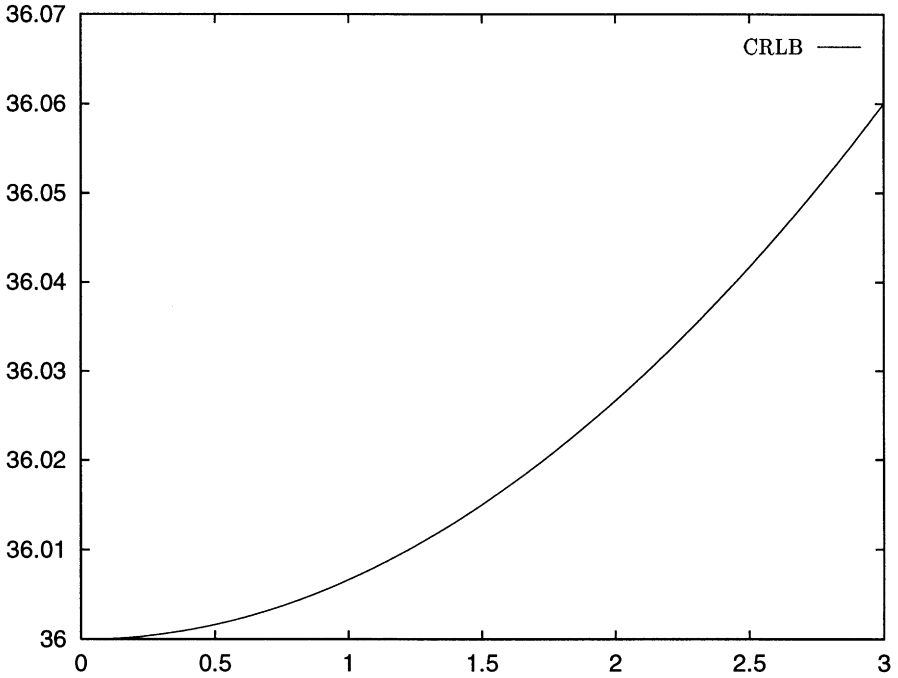

Fig. 2. CRLB of the distance parameter as a function of the width of the intervals, $N = 5000$, $\sigma_0 = 30$, $A_2 = 6$.

value of 31%. Therefore, Eq. (39) agrees well with the experiment. Fig. 4. shows the histogram of the measured B-values.

Fig. 5 shows the histogram of the found estimates of the distance: 3150 estimates out of 10 000 are zero. The ML estimator is clearly biased: not only is there a large amount of zero estimates, but also the non-zero estimates are clearly not centered around the true value.

Fig. 6 shows the probability of resolution, that is, the probability that the factor B is positive, as a function of R. It is important to realize that if the actual distance between the peaks is zero, the probability of resolution is still 50%. This means that the worst case is $P(B > 0) = 0.5$ and the best case $P(B > 0) = 1$. From this figure it can be seen, for instance, that if the resolution factor R, is 1, i.e., if the right-hand side of Eq. (34) equals the left-hand side, the probability of resolution of the ML-estimator is only about 69%. If, on the other hand, a probability of resolution of at least 95% is desired, the ratio $\sigma_0^2/(\sqrt{N}A_2^2)$ should be smaller than $0.3/(2\sqrt{2})$, which is equal to 0.1.

6. Conclusion

Whether two objects, especially Gaussian peaks, can be resolved or not depends on a resolution factor, which is a function of the total number of counts, the distance between the peaks, and the width of the peaks. The ultimate precision with which the locations of the peaks can be estimated and the probability that the maximum likelihood solutions will coincide can be calculated using these quantities. From this, it is possible to deduce rules of thumb for the attainable resolution. For instance, if one desires that the probability of resolving two Gaussian peaks by the Maximum likelihood estimator is at least 95%, the ratio $\sigma_0^2/(NA_2^2)$ must be smaller than 0.1, where σ_0 is the width of the peaks, N the total number of counts and A_2 the distance between the peaks. In case of other demands, other rules can easily be calculated.

In this paper, the focus was on one-dimensional objects. However, further research has shown that extensions to higher dimensions are possible. We

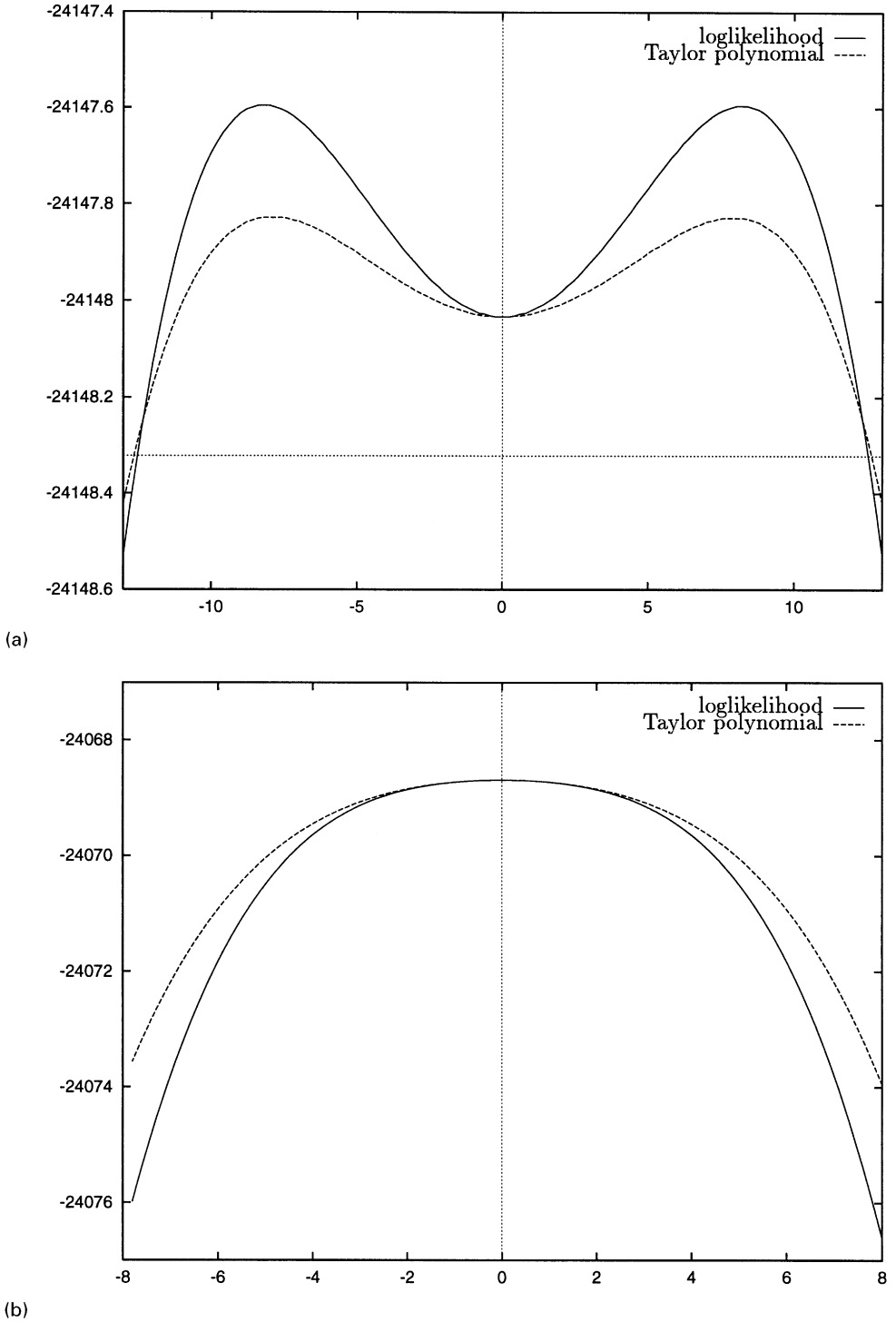

Fig. 3. Two different possibilities for the loglikelihood function and its quadratic Taylor polynomial as a function of the distance: two maxima (a) and one maximum (b).

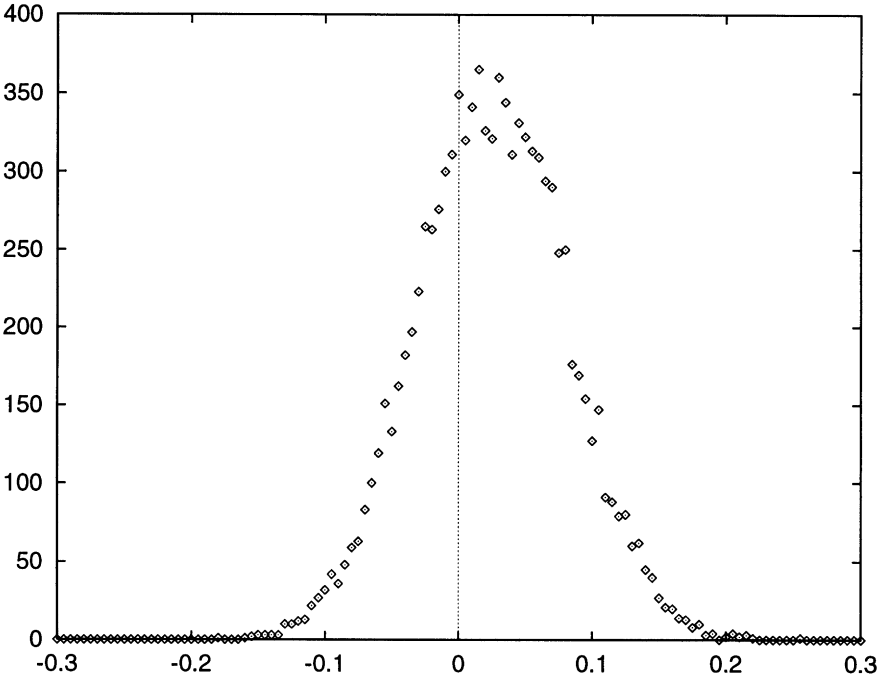

Fig. 4. Histogram of B, $N = 5000$, $\sigma_0 = 30$, $A_2 = 6$.

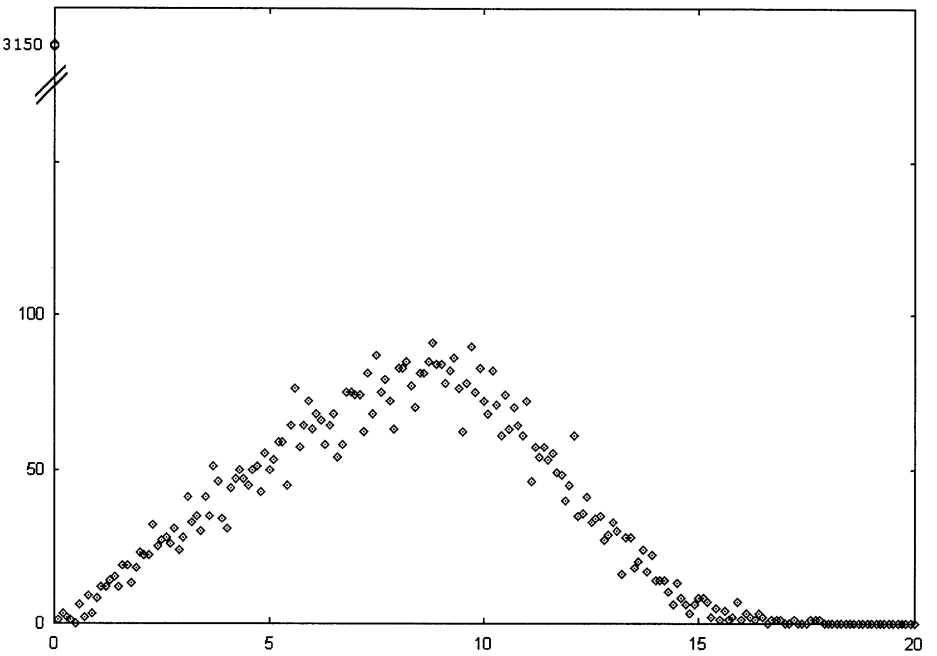

Fig. 5. Histogram of $\hat{A}_2$, $N = 5000$, $\sigma_0 = 30$, $A_2 = 6$.

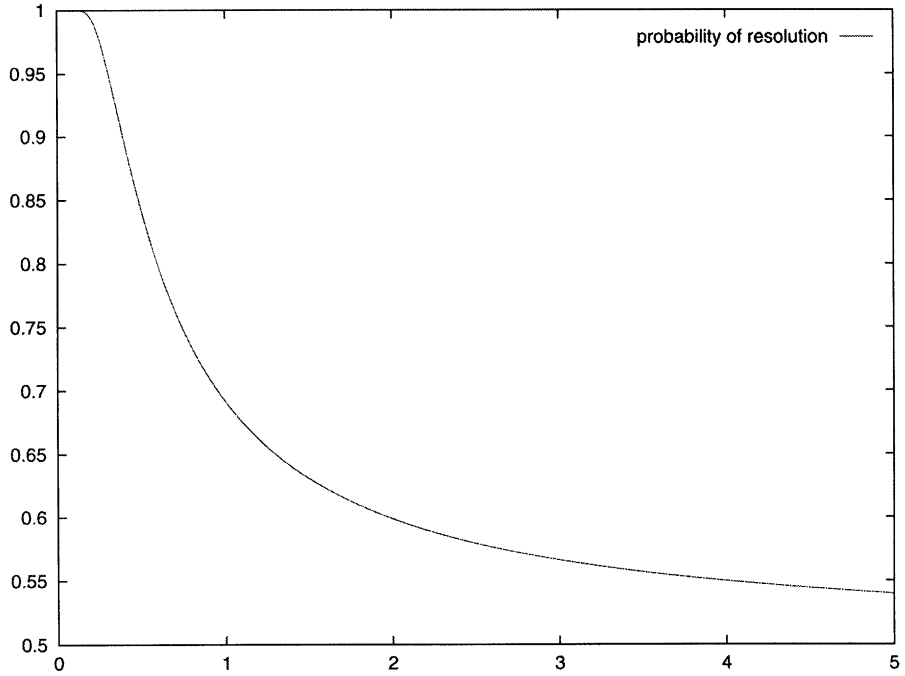

Fig. 6. The probability of resolution, $P(B > 0)$, as a function of $R = 2\sqrt{2}\sigma_0^2/(\sqrt{N}A_2^2)$.

intend to publish results on this subject in the near future. The possibility of extending the results to include coherent object sources will also be studied.

Acknowledgements

E.B. was supported by the following projects of the Antwerp University: *GOA Vision and GOA Generic Optimization*.

References

[1] Lord Rayleigh, Philos. Mag. 47 (1874) 81, 193.
[2] A. Rose, Adv. Electron. Electron Phys. 1 (1948) 131.
[3] A.J. den Dekker, A. van den Bos, J. Opt. Soc. Am. A. 14 (3) (1997) 547.
[4] A. van den Bos, Parameter Estimation, in: P.H. Sydenham (Ed.), Handbook of Measurement Science, vol. 1, wiley, Chichester, 1982, pp. 331–377, chapter 8.
[5] A. Stuart, K. Ord, Kendall's Advanced Theory of Statistics, Edward Arnold, London, 1994.
[6] A. van den Bos, J. Opts. Soc. Am. A. 4 (1987) 1402.
[7] A. van den Bos, Ultramicroscopy 47 (1992) 298.
[8] A. van den Bos, J.H. Swarte, IEEE Trans. Signal Process. 41 (1) (1993) 312.
[9] A. van den Bos, A.J. den Dekker, Ultramicroscopy 60 (1995) 345.
[10] A.J. den Dekker, IEEE Trans. Instr. Measurement 46 (4) (1997) 798.
[11] T. Poston, I.N. Stewart, Catastrophe Theory and its Applications, Pitman, London, 1978.