CORRECTION FOR TRUNCATION OF SIGNALS USING AN ASYMPTOTIC POWER SERIES

Application in X-ray diffraction line profile analysis

ing. A.C. Vermeulen februari 1989

afstudeerverslag

hoogleraar: prof. dr. ir. E.J. Mittemeijer
begeleiders: dr. ir. R. Delhez
dr. ir. Th. H. de Keijser

Technische Universiteit Delft
Faculteit der Scheikundige Technologie en der Materiaalkunde
Vakgroep Fysische en Chemische Metaalkunde
Sectie Struktuur en Fysische Chemie van de Vast Stof

TABLE OF CONTENTS

SUMMARY

A general description has been developed for the tails of signals which extend form $-\infty$ to $+\infty$. To correct for the (inevatable) truncation of measured line profiles the asymptotic power series $\sum_{i=2}^{\infty} \frac{c_i}{x^i}$ (where c_i are constants and x is the distance form the chosen origin) is used. The correction procedure works either in real space or in Fourier space. The latter procedure adds corrections to the (distorted) discrete Fourier transform of the truncated signal yielding a corrected Fourier transform without the characteristic distortions. This makes further analysis of the Fourier transform possible in terms of size (distribution) and strains.

The corrections proposed for truncation takes into account:

(i) wrong estimation of background ("horizontal" truncation)

(ii) lost tails due to truncation ("vertical" truncation).

A computer program has been written for developing and testing the procedure.

The Cauchy, Gaussian and two size-broadened line profiles -signals related to the field of X-ray diffraction- are simulated and used as test cases.

The results are quite satisfactory for the Cauchy line profile and the two size-broadened line profiles. Typical errors of the average column length (crystallite size) of 50–100 % without correction are reduced to 1–10 % after applying the correction. Negative fractions -physically impossible- of column lengths in the column length distribution no longer occur.

1. INTRODUCTION

It is known that truncation of signals heavily distorts their Fourier transforms. The interpretation of these Fourier transforms will yield unreliable data. Unfortunately in many practical cases truncation is unavoidable due to overlap with other signals or the signal continues outside the possible measurement range. Then it is useful to estimate the lost part of the signal or the mutilation of its Fourier transform.

Other ways to treat truncated signals exist, e.g. the use of window functions to smooth the discontinuity of the signal caused by truncation. This kind of smoothing implies changing of the shape of the signal, but it is often applied for want of something better. Fitting some kind of model function through the truncated signal or through its distorted Fourier transform is only applicable if the model function is physically realistic.

The more general method presented here leaves the truncated signal unchanged by predicting only the behaviour of the signal tails lost by truncation. If the Fourier transform exists the ideal signal and its tails can be described by an infinite asymptotic power series derived from the Fourier integral of an ideal signal (see chapter 2).

In practice the observed signal contains the ideal signal with background and noise. Because the observed signal is a sampled signal only a restricted number of terms of the asymptotic power series and the background can be estimated by fitting such series to measured parts of the observed signal tails (see chapter 4).

An estimation of the background can be made and the predicted tails can be added to the truncated signal upto a predetermined range as discussed in section 4.5. This implies that the correction is performed in real space (procedure 1). Alternatively the corrections for the contributions of the predicted tails and for the estimated background can be added to the distorted Fourier transform of the truncated signal (procedure 2) (see chapter 2). The differences between the two procedures mentioned are discussed in detail in section 4.5.

Procedure 1 has a disadvantage: the increase of computer time and memory storage needed for the calculation of the Fourier transform. However, the rather complicated calculation of the corrections of the Fourier transform needed in procedure 2 can be avoided. This procedure has the advantage that the distorted and corrected Fourier transform are both available; a disadvantage is the large sampling distance in Fourier space.

In X-ray diffraction the problem of truncation of line profiles has been discussed and treated by various authors. E.g. Wilkens & Hartmann (1963) estimated the hook effect caused by truncation by assuming a $1/x^2$ behaviour of the profile tails. Young, Gerdes & Wilson (1967) studied the propagation of the truncation error in the Fourier coefficients. Zocchi (1980) suggested the use of the first derivative of the cosine transform modified by an oscillatory factor. Langford (1982) proposed a truncation correction within the variance method of analysis also based on the $1/x^2$ behaviour of the profile tails.

The basic concept of a diffraction line profile and its range have been discussed by Delhez, Keijser, Mittemeijer & Langford (1986). If a truncated line profile is interpreted as a truncated 'component line profile' (as defined in the paper mentioned) procedure 2 can be applied as is shown by examples of simulated line profiles (chapter 6). For a truncated 'total intensity distribution' procedure 1 is possible. In this report only procedure 2 is studied and tested with simulated 'component line profiles'.

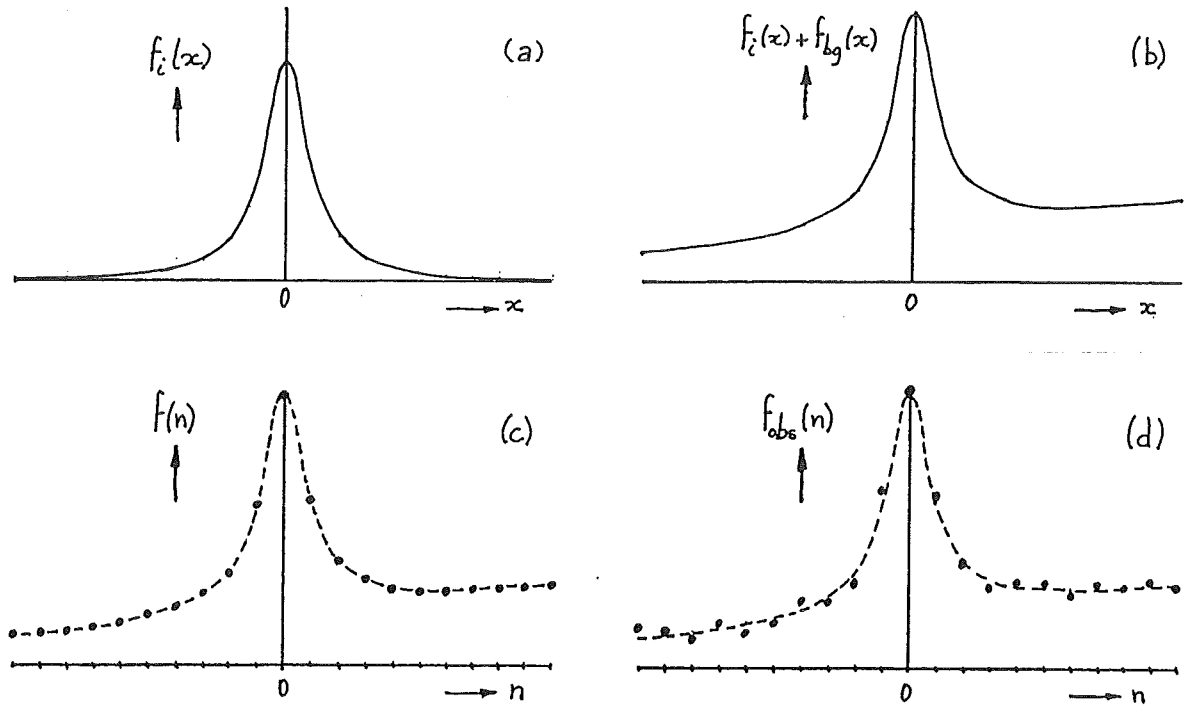

Fig. 2.1 Schematic picture of a) an ideal signal $f_i(x)$ as a continuous function of x b) the ideal signal $f_i(x)$ with background $f_{bg}(x)$ c) the sampled ideal signal with background d) the observed signal as sampled signal which contains the ideal signal with background and noise.

2. MATHEMATICAL BASIS

An ideal signal $f_i(x)$ is defined here as a continuous function of x (see Fig. 2.1a). Besides this ideal signal a continuous background contribution $f_{bg}(x)$ can be present (see Fig. 2.1b). Because of the necessity of numerical calculations and for requirements of the observation technique the ideal signal with background is sampled at usually equidistant values (see Fig. 2.1c). However in practice measurements always include 'noise'. The observed signal $f_{obs}(n)$ (see Fig. 2.1d) is therefore a sampled signal which contains the ideal signal with background and noise.

For the estimation of the effect of truncation of a signal the first condition is a general mathematical description for an ideal signal component including its tails which extend to $-\infty$ and $+\infty$. An example of such a signal component I_{com} as a function of h_3 is given in Fig.2.2. The treatment given below does not imply the condition that the signal has only positive ordinate values.

The Fourier transform pair belonging to the signal component is given by

$$I_{com}(h_3\text{-}l) = K \int_{-\infty}^{\infty} F_{com}(t)\, \exp[2\pi i t(h_3\text{-}l)]\, dt \qquad (1a)$$

$$F_{com}(t) = \frac{1}{K} \int_{-\infty}^{\infty} I_{com}(h_3\text{-}l)\, \exp[-2\pi i t(h_3\text{-}l)]\, d(h_3\text{-}l) \qquad (1b)$$

where $I_{com}(h_3\text{-}l)$ = intensity of the signal as a function of the dimensionless variable h_3 with reference point l, $F_{com}(t)$ = Fourier transform of signal component, t = dimensionless Fourier space variable and K is a proportionally constant.

Real Space

When $I_{com}(h_3\text{-}l)$ is a real function and if the Fourier transform $F_{com}(t)$ exists Eq.(1a) can be written as :

$$I_{com}(h_3\text{-}l) = 2K \int_{0}^{\infty} \frac{F_{com}(t)}{2\pi i(h_3\text{-}l)}\, d\, \exp[2\pi i t(h_3\text{-}l)] \qquad (2)$$

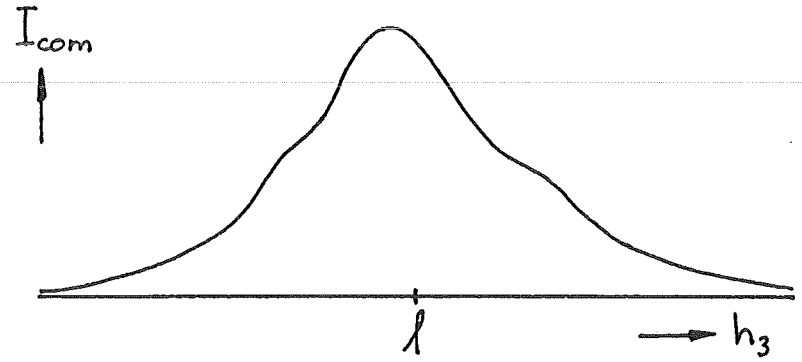

Fig. 2.2 Schematic picture of an ideal signal component I_{com} as a function of a dimensionless variable h_3 with reference point l. No background present.

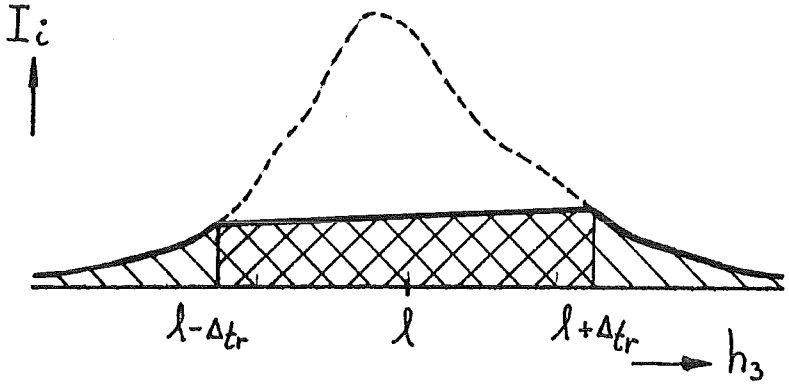

Fig. 2.3 Schematic picture of the ideal signal I_i. The lost part of the signal I_{lost} is shaded. Δ_{tr} is half of the truncation range.

With

$$F_{com}(t) = A_{com}(t) + iB_{com}(t) \qquad (3)$$

and the conditions that $A_{com}(t)$, $B_{com}(t)$ and all their derivatives $A^{(n)}(t)$ and $B^{(n)}(t)$ are zero for $|t| \rightarrow \infty$ the following asymptotic power series can be derived for $|h_3-l|>0$ by repeated integration by parts (see Erdélyi, 1956 for an extensive discussion):

$$I_{com}(h_3-l) = 2K\left[\frac{-A^{(1)}(0)}{(2\pi)^2(h_3-l)^2} + \frac{A^{(3)}(0)}{(2\pi)^4(h_3-l)^4} - \frac{A^{(5)}(0)}{(2\pi)^6(h_3-l)^6} + ..\right]$$

$$+2K\left[\frac{B^{(2)}(0)}{(2\pi)^3(h_3-l)^3} - \frac{B^{(4)}(0)}{(2\pi)^5(h_3-l)^5} + ..\right]. \qquad (4)$$

Cheary & Grimes (1972) already used this type of formula in the field of X-ray diffraction.

In compact notation:

$$I_{com}(h_3-l) = \sum_{k=1}^{\infty} \frac{c_{2k}}{(h_3-l)^{2k}} + \sum_{m=1}^{\infty} \frac{c_{2m+1}}{(h_3-l)^{2m+1}} \qquad (5a)$$

where

$$c_{2k} = 2K(-1)^k \frac{A^{(2k-1)}(0)}{(2\pi)^{2k}} \qquad ,k = 1,2,3,... \qquad (5b)$$

and

$$c_{2m+1} = 2K(-1)^{m+1} \frac{B^{(2m)}(0)}{(2\pi)^{2m+1}} \qquad ,m = 1,2,3,... \qquad (5c)$$

The even coefficients c_{2k} describe the symmetric part and the odd coefficients c_{2m+1} describe the asymmetric part of the signal component with respect to reference point l.

Truncation of the signal component and removal of estimated background will cause a loss of a part of the signal (see Fig 2.3). The truncation points $l+\Delta_{tr}$ and $l-\Delta_{tr}$ are chosen symmetrically with respect to the reference point l.

Because Eqs.(5abc) describe the complete signal component they can also be used for the description of the tails of the signal lost by truncation:

$$I_{lost}(h_3-l)= \sum_{k=1}^{\infty} \frac{c_{2k}}{(h_3-l)^{2k}} + \sum_{m=1}^{\infty} \frac{c_{2m+1}}{(h_3-l)^{2m+1}} \quad , |h_3- l| \geq \Delta_{tr} \qquad (6a)$$

with $2\,\Delta_{tr}$ = truncation range.

In practice usually background is estimated by interpolating linearly between the intensities at both truncation points. This overestimates the background compared to the estimate using the procedure proposed here by the cross-hatched area of Fig 2.3. This can be described with:

$$I_{lost}(h_3-l)= \sum_{k=1}^{\infty} \frac{c_{2k}}{\Delta_{tr}^{2k}} + (h_3-l) \sum_{m=1}^{\infty} \frac{c_{2m+1}}{\Delta_{tr}^{2m+1}} \quad , |h_3-l| \leq \Delta_{tr} \qquad (6b)$$

Fourier space

By calculating the Fourier transform of Eqs. (6ab) a general formula for the Fourier transform of the lost parts of the signal component for discrete values of t and for integer values of t' is obtained (see Appendix Q):

$$A_{lost}(t) = \quad 0 \quad + \frac{2}{K} \sum_{k=1}^{\infty} \frac{c_{2k}}{\Delta_{tr}^{2k-1}} \alpha_k(t') \qquad (7a)$$

$$A_{lost}(0) = \frac{2}{K} \sum_{k=1}^{\infty} \frac{c_{2k}}{\Delta_{tr}^{2k-1}} + \frac{2}{K} \sum_{k=1}^{\infty} \frac{1}{2k-1} \frac{c_{2k}}{\Delta_{tr}^{2k-1}} \qquad (7b)$$

$$B_{lost}(t) = -\frac{2}{K} \frac{(-1)^{t'+1}}{\pi t'} \sum_{m=1}^{\infty} \frac{c_{2m+1}}{\Delta_{tr}^{2m}} - \frac{2}{K} (\pi t') \sum_{m=1}^{\infty} \frac{c_{2m+1}}{\Delta_{tr}^{2m}} \frac{\alpha_m(t')}{2m} \qquad (7c)$$

$$B_{lost}(0) = \quad 0 \quad + \quad 0 \qquad (7d)$$

where $t' = 2\Delta_{tr}.\,t$ and $\alpha_k(t')$ or $\alpha_m(t')$ are tabulated functions as defined in Appendix R and depend only on k or m and t', not on Δ_{tr}.

The first terms on the right-hand side of Eqs. (7abcd) are the contributions of the part of the signal component lost due to wrong estimation of background and the second terms are the contributions of the lost tails of the signal. In practice we have to change from the continuous Fourier transform to the discrete Fourier transform (DFT), because of the necessity of numerical calculations.Then the following relationship holds:

$$t = \frac{t'}{2\Delta_{tr}} \tag{8}$$

Here t' takes integer values only. Eq.(8) implies that the range of the discrete Fourier transformation is taken equal to the truncation range $2\Delta tr$.

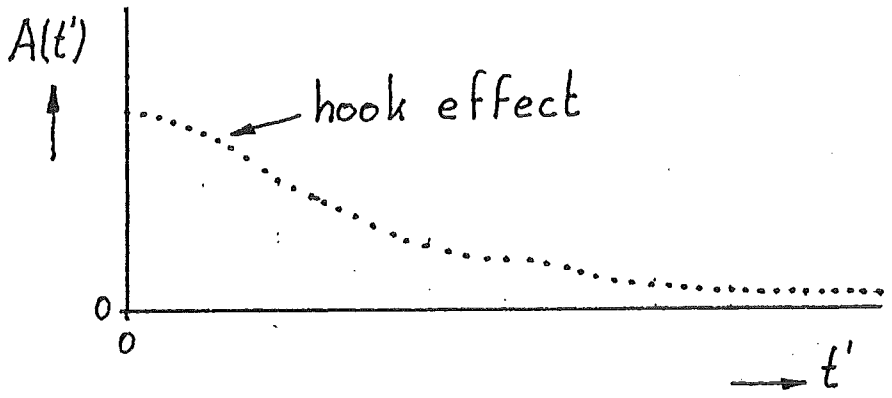

Fig. 3.1 Schematic picture of normalized Fourier coefficient curve with a negative initial curvature (hook effect).

3. EFFECTS OF TRUNCATION

In the following the effects of truncation will be discussed for the field of X-ray diffraction pattern analysis. It is obvious that the truncation problem occurs in many research field and it is in particular important when using Fourier analysis.

In X-ray diffraction the words 'line profile' or 'profile' are commonly used instead of the word 'signal'. The symbols used in chapter 2 have a more specific meaning in X-ray diffraction. The function $I_{com}(h_3\text{-}l)$ is a component of the total intensity distribution in reciprocal space along a [00l] direction, where the dimensionless variable h_3 is defined obeying

$$h_3 = \frac{2a_3 \sin\theta}{\lambda} \qquad (9)$$

with a_3 = the lattice parameter in the [00l] direction, θ = half of diffraction angle 2θ and λ = X-ray wavelength (Delhez, Keijser, Mittemeijer & Langford, 1986).

Truncation introduces errors in profile characteristics including the Fourier coefficients. These errors are propagated to the size- and strain parameters derived from the Fourier coefficients (e.g. Delhez, Keijser, Mittemeijer & Langford, 1988).

In experimental Fourier coefficient curves often a negative initial curvature is observed (see Fig 3.1). This so-called hook effect causes a serious error in average column length and physically impossible negative fractions of column length in the column length distribution occur (e.g. Young, Gerdes & Wilson, 1967).

Truncation is not the only possible cause of the hook effect, it can be also due to strain present in the sample. Existing simple correction methods for the observed hook effect are therefore dangerous (Delhez, Keijser & Mittemeijer, 1982).
For the case of pure size broadening a $1/(h_3\text{-}l)^2$ behaviour of the profile tails can be deduced. Assuming this behaviour Wilkens & Hartmann (1963) estimated the hook effect caused by truncation and found a good agreement between observed and calculated hook effect in the case

where size broadening is dominant, but the agreement was less satisfactory in the case of both size and strain broadening.

Delhez, Keijser, Mittemeijer & Langford (1986) defined a 'total intensity distribution' and a 'component line profile' with respectively $1/\sin^2\pi h_3$ and $1/(h_3-l)^2$ behaviour of the profile tails for the case of pure size broadening (Appendix T). A description of the effect of truncation on the determination of the average column length is deduced in the paper mentioned.
However, the $1/(h_3-l)^2$ behaviour of the tails in the 'component line profile' approach is only the first term of the asymtotic power series Eq. (4). Probably a similar statement can be made for the $1/\sin^2\pi h_3$ behaviour of the tails in the 'total intensity distribution' approach, but a direct derivation of the analogue of the asymptotic power series has not been found yet.

Related to the truncation problem in X-ray diffraction pattern analysis the following new ideas are presented:

- Eq.(4) holds for the complete profile and therefore for both tails at the same time
- describe the profile tails with as many terms of Eq. (4) as can be calculated with sufficient accuracy
- the odd terms enable a description of the asymmetric part of the profile
- the origin can in principle be chosen arbitrarily
- the limitation to profiles with only size broadening can be dropped.

4. CORRECTION FOR TRUNCATION IN PRACTICE

4.1 *Estimation of coefficients* c_{2k} *and* c_{2m+1}

In chapter 2 the lost parts of the ideal signal (Eqs. 6ab) and the mutilation of its discrete Fourier transforms (Eqs. 7abcd) are described. These descriptions are exact when all coefficients c_{2k} and c_{2m+1} (k, m= 1, 2, 3, $\ldots\infty$) of Eq. (5a) are known. In principle the information for determining these coefficients is included in the truncated part of the ideal signal. However, in practice the observed signal contains the ideal signal with background and noise. Because the observed signal is a sampled signal only a limited number of coefficients c_{2k} and c_{2m+1} can be estimated with sufficient accuracy.

In this investigation a least-squares procedure -using the χ^2 figure-of-merite function- is applied to determine these coefficients. χ^2 is defined in Appendix K: it is the sum of the ratios of the squared differences between measured and fitted intensities and the squared standard deviations of the measured intensities.

In X-ray diffractometry the standard deviation in the observed intensities [illegible] of the number of counts (provided

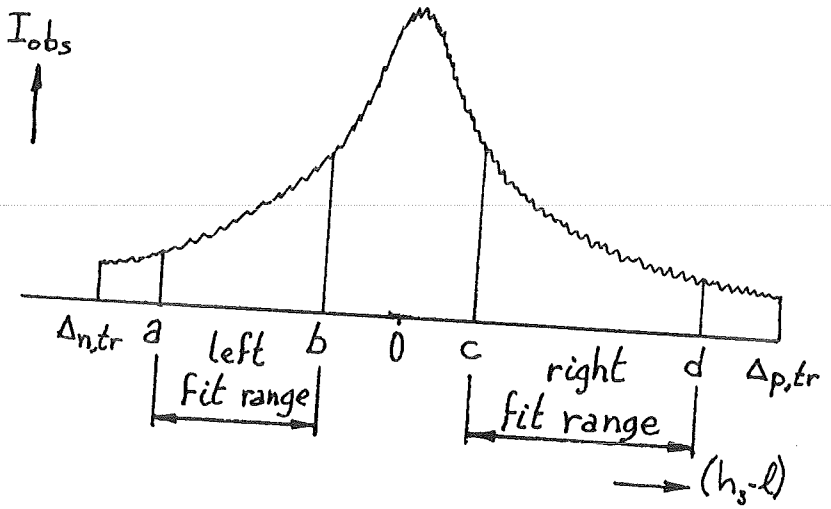

Fig. 4.1 Schematic picture of the measured parts a-b and c-d of the signal tails used for fitting.

$\Delta_{n,tr}$ = truncation point at negative side of signal

$\Delta_{p,tr}$ = truncation point at positive side of signal.

Method 1 *Direct fitting to the observed signal*
A restricted number of terms of the asymptotic series (5a) and the background are estimated with a least-squares fitting procedure by fitting:

$$I_{fit}(h_3\text{-}l) = c_o + c_1(h_3\text{-}l) + \sum_{k=1}^{N} \frac{c_{2k}}{(h_3\text{-}l)^{2k}} + \sum_{m=1}^{M} \frac{c_{2m+1}}{(h_3\text{-}l)^{2m+1}} \quad (10)$$

to measured parts at the signal tails (see Fig. 4.1).
Because Eq. (5a) holds for the ideal signal as a whole, Eq. (10) can be fitted to both sides of the observed signal simultaneously.

Method 2 *Variation of truncation range*
From the observed signal data the integral intensity can be calculated as a function of the truncation range.
The integral intensity of the observed signal contains a background contribution of $2c_o\Delta'_{tr}$ when assuming the background to be linear (see Appendix S). A limited number of terms of the asymptotic series (7b) with the background contribution are estimated with a least-squares fitting procedure by fitting:

$$A^*_{obs}(\Delta'_{tr}) = A^*_i - 2\sum_{k=1}^{N} \frac{1}{2k-1} \frac{c_{2k}}{(\Delta'_{tr})^{2k-1}} + 2c_o\Delta'_{tr} \quad (11)$$

to the sampled integral intensity $A^*_{obs}(\Delta'_{tr})$ (see Fig. 4.2). This formula is deduced in Appendix S. With Eq. (11) only the even coefficients c_{2k} and c_o can be estimated. The odd coefficients c_{2m+1} and c_1 can be estimated with an analogous formula for the first moment (see Appendix S).

Also the maximum intensity or the integral breadth can be used to determine the coefficients c_{2k} and c_{2m+1} (see Appendix S for references). This method is closely related to the truncation correction of the well-known variance method of analysis in X-ray powder diffraction (Wilson, 1962). The truncation correction as discussed by Langford (1982) contains only one term of the asymptotic series and is applied in combination with removal of the estimated background. Eq. (11) shows that the preceding removal of the background is not necessary.

Obviously both methods are equivalent, because the same observed data points are used for determining the coefficients c_{2k} and c_{2m+1}. With method 1 both symmetry and asymmetry at the signal and further background level and slope can be estimated by fitting just one equation. With method 2 symmetry and asymmetry are seperated. The even coefficients c_{2k} and c_0 have to be estimated with e.g the integral intensity and the odd coefficients c_{2m+1} and c_1 with e.g. the centroid. Disadvantage of the fitting procedure of both methods is that the basis functions are non-orthogonal i.e. the basis functions are too much alike. Variance and covariance of the estimated coefficients will increase drastically with increasing number of coefficients to be estimated. The set of even basis functions is however orthogonal with respect to the set of odd functions. So symmetry and asymmetry contributions are well separable. Other factors limiting the number of coefficients estimated will be discussed in the next section.

In the following the direct fitting method (method 1) is adopted as standard procedure for estimation of background and coefficients c_{2k} and c_{2m+1}. The truncation variation method (method 2) will not be treated further.

4.2. *Maximum number of estimated coefficients c_{2k} and c_{2m+1}*

In least-squares fitting the number of estimated parameters must always be less than the number of data points. So the maximum number of coefficients c_{2k} and c_{2m+1} of Eq. (5a) which can be estimated depends on the number of data points present in the fitranges (see Fig. 4.1). It is clear that a large truncation range and a small sampling distance both increase the number of data points. The fitranges should be chosen as large as possible. The outer boundaries of the fitranges a and d (see Fig. 4.1) can be taken equal to the truncation points $-\Delta_{n,tr}$ and $\Delta_{p,tr}$. The inner boundaries b and c (see Fig. 4.1) should be optimized, because more terms of the asymptotic series are needed for an accurate description (Cf. Appendix U) of data points near the origin where the asymptotic series diverges.
Increasing the accuracy of the observed data points also increases the number of parameters that can be calculated with sufficient accuracy. In

X-ray diffraction a higher accuracy can be achieved by increasing the counting time. Because in practice the total measurement time is limited a special counting strategy (e.g. longer counting time for the higher intensities) has to be adopted. This will not be discussed further.

4.3. *Choice of origin and truncation points*

A symmetric truncation with respect to the origin must be achieved, because of the required calculation of the discrete Fourier transform and the correction for truncation. For optimum use of the measured data the origin should be chosen at the middle of the measurement range. However, this will generally not be optimum choice for the least-squares fitting procedure, because it can be shown that a significant change in the position of the origin leads to a change in the number of terms of the asymptotic series (5a) needed for a sufficiently accurate description of the signal tails (see Appendix U). Therefore the choice of the origin should be optimized within the least-squares fitting procedure. After finding the optimum origin and the estimation of the coefficients c_{2k} and c_{2m+1} the truncation points are generally asymmetric with respect to the origin (see Fig. 4.3). Because of the restriction of symmetric truncation for calculating the discrete Fourier transform and the correction for truncation, new truncation points Δ_{tr} and $-\Delta_{tr}$ must be chosen. We have two possibilities (Cf. Fig. 4.3).

Possibility 1 *Choice of Δ_{tr} equal to the length of the shorter tail*

This implies further truncation causing a loss of information.

Possibility 2 *Choice of Δ_{tr} equal to the length of the longer tail.*

This implies an extrapolation with Eq. (10).

Possibility 1 is not to be preferred because of the loss of information. This will result in a sampling distance in Fourier space larger than for possibility 2. Therefore possibility 2 is recommended.

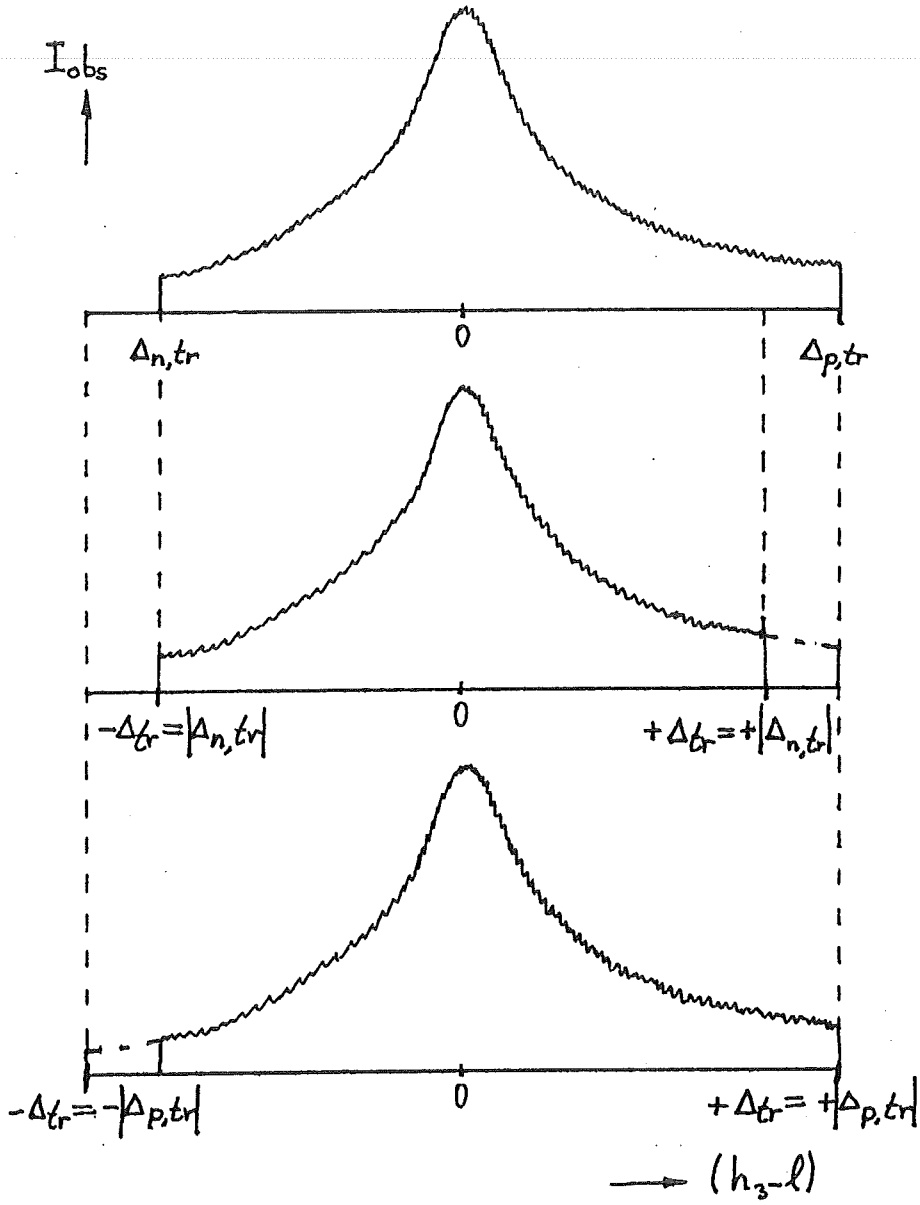

Fig. 4.3 Schematic picture of asymmetric truncation
a) with truncation points $\Delta_{n,tr}$ and $\Delta_{p,tr}$ where $|\Delta_{n,tr}| < |\Delta_{p,tr}|$.
b) Δ_{tr} chosen equal to $\Delta_{n,tr}$ implying further truncation
c) Δ_{tr} chosen equal to $\Delta_{p,tr}$ implying extrapolation.

4.4 *Removal of background*

Before evaluation of the signal, e.g. calculation of the Fourier transform, the estimated background has to be removed. If one of the methods described in section 4.1 is applied the background $c_0 + c_1 (h_3\text{-}l)$ is removed without causing "horizontal" truncation i.e. overestimation of the occurring background as discussed in chapter 2.
Therefore when applying Eqs (7abcd) for correcting the discrete Fourier transform for lost tails only the second terms on the right-hand side have to be used. It is remarked that the first terms on the right-hand side of Eqs. (7abcd) can be used for a "negative correction" to approximate the situation of both overestimated background and lost tails due to truncation (Vermeulen, 1988-0711a).

4.5 *Calculation of the corrected DFT*

Two different procedures for calculating the corrected discrete Fourier transform (DFT) will be discussed. For both procedures the estimated background has to be removed after the tail fitting procedure as discussed in section 4.4.

Procedure 1 *Adding tails*
The predicted tails are added to the truncated signal upto the predetermined range $(l\text{-}1/2, l\text{+}1/2)$ as discussed for X-ray profile analysis in Appendix T. The range of the corrected signal will be larger by a factor 5-50 than the range of the truncated signal. This will result in a DFT with a smaller sampling distance in Fourier space and with less distortion than for the DFT of the truncated signal (Vermeulen, 1988-0726a, 0923a, Appendix W).
However, because of the long tails added a considerable increase in the number of data occurs leading to an increase in computation time and memory usage needed for calculation and storage of the DFT. This aspect has not been studied yet.

Procedure 2 *Correction of the distorted DFT*
The discrete Fourier transform of the truncated signal is calculated and corrected for lost tails. The mathematics of the correction for lost tails is discussed in chapter 2: the second terms on the right-hand side of the

Eqs. (7abcd) can be used for correcting the distorted DFT (see also section 4.4). Because the range of the signal is unchanged this will result in a DFT with a relatively large sampling distance in Fourier space but with only a little distortion remaining (Vermeulen, 1988-0726a, 0923a, Appendix W). This procedure is correct for signals with tails extending to $-\infty$ and $+\infty$. The computer time and memory usage needed for calculating the corrections are negligible provided that a table with the α-functions is available. These functions are very difficult to calculate (see Appendix R).

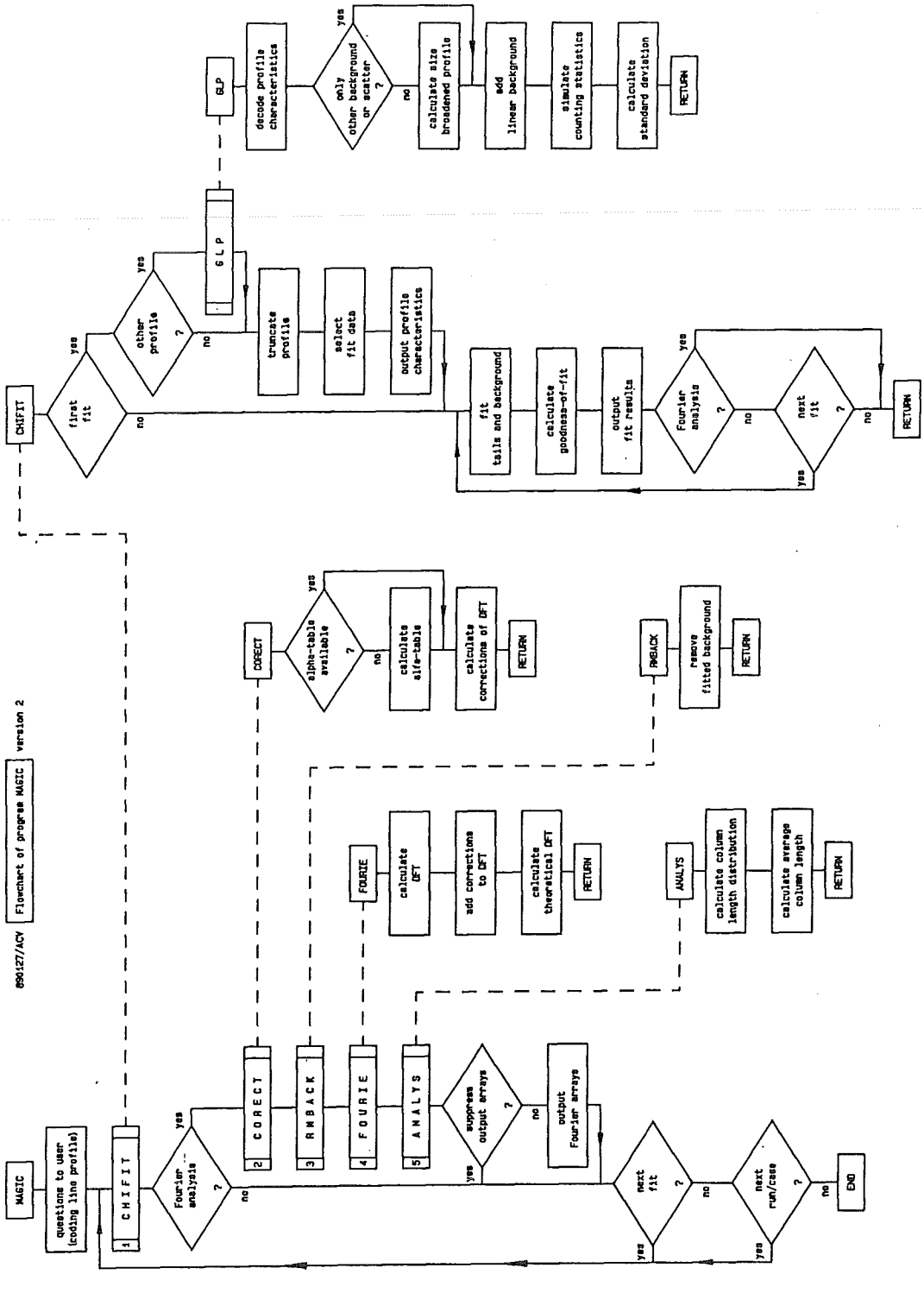
MAGIC
questions to user (coding line profile)
CHIFIT
Fourier analysis ?
yes
no
CORECT
RMBACK
FOURIE
ANALYS
suppress output arrays ?
output Fourier arrays
next fit ?
next run/case ?
END
CHIFIT
first fit ?
other profile ?
GLP
truncate profile
select fit data
output profile characteristics
fit tails and background
calculate goodness-of-fit
output fit results
Fourier analysis ?
next fit ?
RETURN
GLP
decode profile characteristics
only other background or scatter ?
calculate size broadened profile
add linear background
simulate counting statistics
calculate standard deviation
RETURN
CORECT
alpha-table available ?
calculate alfa-table
calculate corrections of DFT
RETURN
RMBACK
remove fitted background
RETURN
FOURIE
calculate DFT
add corrections to DFT
calculate theoretical DFT
RETURN
ANALYS
calculate column length distribution
calculate average column length
RETURN
890127/ACV
Flowchart of program MAGIC
version 2

Fig. 5.1

5. BRIEF DESCRIPTION OF THE PROGRAM

For the application of the method explained in the previous chapters a computer program -written in Fortran-77 and provisionally called MAGIC- has been developed. The five main stages of the program will be discussed shortly (see Fig. 5.1).

Stage 1 *Generation of simulated diffraction data and determination of the lost part of the profile.*

In accordance with the previous chapters all available profiles are of the 'component line profile' type which implies that a profile extends form $-\infty$ to $+\infty$ (Cf. Appendix T). The profile functions used describe the size broadening in X-ray diffraction. A linear background is added to the calculated profile and the effect of counting statistics is simulated. Details of the calculation of the profile data, background and counting statistics are discussed in Appendix G.

Then an origin for the tail fitting procedure is chosen and the generated profile is truncated from $-\Delta_{tr}$ to Δ_{tr} (see section 4.3). The origin chosen automatically becomes the origin for the discrete Fourier transformation and $2\Delta_{tr}$ becomes the range of the discrete Fourier transformation. Moreover symmetical truncation is required for the application of Eqs. 7abcd to correct the DFT for truncation.

For simplicity of programming the fitting ranges at the extremeties of the simulated profile are chosen to be of equal length and the ranges are therefore symmetically with respect to the origin chosen.

With the χ^2 least-squares fitting procedure discussed in section 4.1 Eq. (10) is fitted to the intensity data in the ranges selected. This implies the calculation of a goodness-of-fit parameter along with the estimation of the coefficients c_{2k} and c_{2m+1} and background level and slope (Cf. stage 3). Standard deviations of the intensity data are used for the calculation of χ^2 and of the goodness-of-fit (Cf. Appendix K). These standard deviations are estimated as the square root of the number of counts (Cf. section 4.1).

Stage 2 *Calculation of the corrections for the lost part of the profile to be added to the DFT.*

The corrections of the DFT for the lost part of the profile are calculated according to Eqs. 7abcd using the α-functions as defined in Appendix R. Because these functions are problem-independent (i.e. independent of the truncation point Δ_{tr}) they have to be calculated only once.
The corrections are split into two contributions (i) the corrections of the DFT for the lost part of the profile due to overestimation of the background ("horizontal" truncation) and (ii) the corrections for tails lost due to truncation ("vertical" truncation). These contributions are described by the first and second terms on the right-hand sides of Eqs. (7abcd) respectively.

Stage 3 *Removal of the background*

The estimation of the background is made in the fitting procedure of stage 1. This way of removing the background implies that no overestimation of the background ("horizontal" truncation) occurs (Cf. section 4.4).

Stage 4 *Discrete Fourier transformation of the truncated profile and calculation of the corrected DFT.*

For the fast calculation of the DFT a so-called mixed-radix routine is used (Singleton, 1969). The calculated DFT (i.e. of the only vertically truncated profile) is corrected by adding the corrections calculated in stage 2. Only the corrections for tails lost due to truncation ("vertical" truncation) have to be added (see stage 3).

It is of interest to compare e.g. the average column length obtained from
(a) the DFT calculated using the present method
(b) the analytical Fourier transform
(c) the DFT calculated with the usual way of removing the background (i.e. subtracting of a straight line background between the profile extremeties).
The values of the analytical Fourier transform of the simulated line profile are calculated at the same t'-values as dictated by the calculated DFT (see Appendix G).

In case (c) the background is taken too high ("horizontal" truncation). The contribution to the DFT caused by this overestimation of the background (cross-hatched area in Fig. 2.3) is calculated in stage 2. By subtracting this contribution from the DFT of the only vertically truncated profile the DFT of case (c) can be estimated (Cf. Eq. 7 and the text below it).

Stage 5 *Quantities obtained from the DFT's*

The three DFT's mentioned above and their moduli are used for the calculation of the average column length and the column length distribution as discussed in Appendix H.

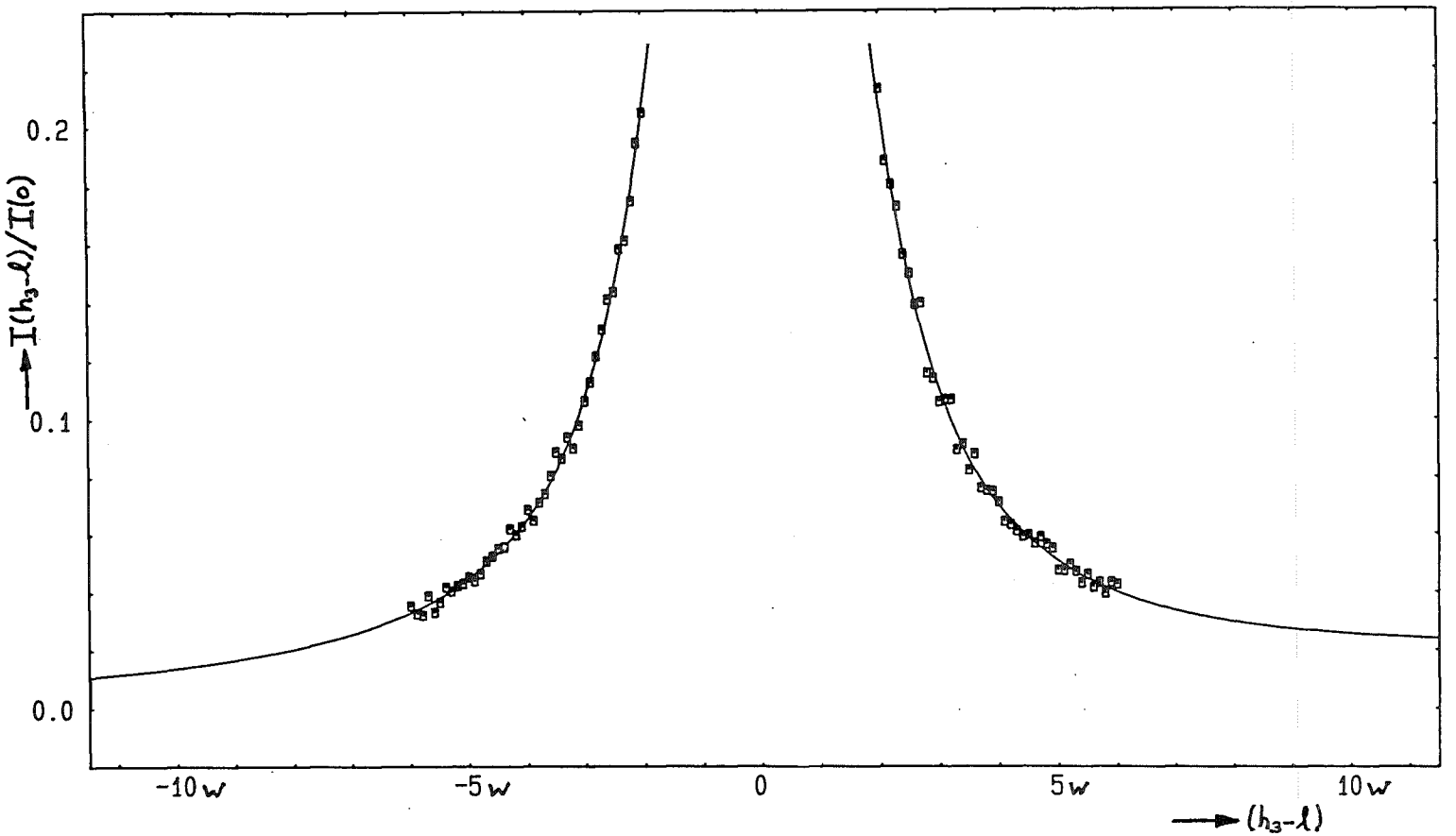

Fig. 6.1 Typical example of the simulated intensity data (squares) and fitted tails (solid line). Only the intensity data in the fit ranges (only in this figure: $|2w| \leq (h_3-\ell) \leq |6w|$) are shown. Cauchy line profile with linear background. Fitted numbers of even terms (N) and odd terms (M) of the asymptotic series: (N, M) = 2,1.

6. TEST OF THE PROCEDURE

The method as described in the previous chapters is tested with the program MAGIC in three steps. First it is tested if the χ^2 fitting procedure yields the correct parameter values with simple profiles for which the asymptotic series (c_{2k}-values) and the Fourier transform are known analytically. The next step is to test the method with complex profiles for which the asymptotic series (c_{2k}-values) is not known, but the Fourier transform and physical characteristics are calculable. In this step also the effect of counting statistics is studied using 10 different simulations. The third step tests the use of the goodness-of-fit parameters Q ("average" misfit) and R ("systematic misfit") for the optimization of the fit procedure to find the best model (i.e. the appropriate number of terms of the asymptotic series), the best origin position and the 'best' inner boundary of the fit range.

The parameter values given in Table 1 are used for all simulated intensity data in the experiments discussed below. For a typical example of the simulated intensity data and fitted tails see Fig. 6.1. In all experiments the background $c_0 + c_1 (h_3 - l)$ is fitted simultaneously with the asymptotic series.

Intensity	counts at top		10000
profile	half width at half maximum (w)		1.0 w
$I(h_3 - l)$	sampling distance		0.1 w
Background	level [counts at top]	c_0	100
$c_0 + c_1(h3-l)$	slope [counts/w]	c_1	2.0

Table 1 Parameters values applied for simulated intensity data. These values are used for all profile functions: Cauchy, Gaussian and the two size-broadened profiles (Cf. Appendix G).

Step 1 *Test with profiles for which the asymptotic series and the Fourier transform are known analytically*

A Cauchy function line profile cannot be interpreted as a size-broadened profile, because the physical interpretation of the column length distribution function p_{N3} of these profiles is not clear. However, a Cauchy profile provides a very useful test, because the DFT, the average area-weighted column length $<N_3>_a$, the column length distribution function p_{N3} and the coefficients c_{2k} of the asymptotic series, are all known. Numerical data for the simulations performed are explained in Appendix G.

The simulations show that for large $(h_3\text{-}l)$-values only the first term of the asymptotic series (5a) will describe the tails of the profile sufficiently accurate (see $|\Delta_{tr}| = 5.0 \ldots 15.0$ in Table 2) as expected. More terms are needed for lower $(h_3\text{-}l)$-values (see Table 3) because (i) in this (Cauchy) case the $1/(h_3\text{-}l)^2$ behaviour suffices only for fitting ranges impractically far from the origin and (ii) the counting statistics will mask the behaviour of the tails (see $|\Delta_{tr}| = 17.5$ and 20.0 in Table 2).

$\|\Delta_{tr}\|$	5.0 w	7.5 w	10.0 w	12.5 w	15.0 w	17.5 w	20.0 w
c_2	7292	9267	9199	9917	10778	9167	12452
Δc_2	48	200	475	928	1595	2512	3713
$<N_3>_{a,cor}$	149.8	125.1	120.1	113.2	106.3	112.4	96.6
$<N_3>_{a,th}$	134.7	122.4	116.5	113.1	106.3	109.2	108.1

Table 2 Fit results of Cauchy line profile with different truncation point $|\Delta_{tr}|$ and with fixed length of fit ranges (3.5 w) at both sides of the profile. The theoretical value of coefficient c_2 is 10000. Δc_2 is the standard deviation of fitted coefficient c_2. Only the first term of the asymptotic series and a linear background is fitted (N,M = 1,0; Cf Table 3). $<N_3>_{a,cor}$ is the average column length as calculated with the corrected DFT, $<N_3>_{a,th}$ is the average column length as calculated with the theoretical Fourier transform with the same sampling distance as the DFT.

If the fit procedure gives accurate values of the c_{2k}-values the tails are described correctly and the corrected DFT will yield reliable values for the average column length $\langle N_3 \rangle_{a,cor}$ (see Table 3). For this (Cauchy) case the accuracy of the coefficient c_2 is of most importance and can only be reached when more terms of the asymptotic series are fitted.
The order of magnitude of the corrections for truncation is (largest first) (i) correction for wrong estimation of background ("horizontal" truncation) (ii) correction for lost tails due to truncation ("vertical" truncation) with only $c_2/(h_3\text{-}l)^2$ (iii) correction for lost tails with higher order terms of the asymptotic series. This is shown by the average column lengths $\langle N_3 \rangle_{a,\ h\ +\ v}$, $\langle N_3 \rangle_{a,v}$ and $\langle N_3 \rangle_{a,cor}$ in comparison with $\langle N_3 \rangle_{a,th}$ in Table 3. The value of $\langle N_3 \rangle_{a,th}$ is not equal to $\langle N_3 \rangle_a$ of the ideal profile (100 for this Cauchy case), because of the relatively coarse sampling in Fourier space.

(N,M)	(1,0)	(2,1)	(3,2)	(4,3)	(5,4)
c_0	232.6	116.1	89.8	74.3	78.5
c_2	7292	9421	10100	10614	10443
c_4	–	-5791	-9982	-15016	-12667
c_6	–	–	6697	24532	10986
c_8	–	–	–	-20308	13621
c_{10}	–	–	–	–	-30416
Δc_2	48	160	423	1021	2373
$\langle N_3 \rangle_{a,h+v}$	243.0	231.6	230.2	229.7	229.7
$\langle N_3 \rangle_{a,v}$	184.0	170.0	167.2	165.7	166.1
$\langle N_3 \rangle_{a,cor}$	149.8	136.7	134.0	132.2	132.7
$\langle N_3 \rangle_{a,th}$			134.7		

Table 3 Fit results of Cauchy line profile with different numbers of even terms (N) and odd terms (M) of the asymptotic series. Truncation point Δ_{tr} and inner boundary of fit range are fixed at 5.0 w and 1.5 w respectively. The theoretical c_{2k}-values are: $c_0 = 100$, $c_2 = 10000$, $c_4 = -10000$, $c_6 = 10000$, $c_8 = -10000$ and $c_{10} = 10000$. $\langle N_3 \rangle_{a,v+h}$ is the average column length of the "horizontal" and "vertical" truncated profile, $\langle N_3 \rangle_{a,v}$ is the average column length of the only "vertical" truncated profile.

The Gaussian line profile cannot be interpreted as a size-broadened profile, because the column length distribution function p_{N3} does not exist. As a consequence, the average column length $\langle N_3 \rangle_a$ does not exist either. Here $\langle N_3 \rangle_a$ is only used as a scaling factor, because $\langle N_3 \rangle_a$ is related to the breadth of the profile. Because theoretically all coefficients are zero, problems are expected in fitting the asymptotic series. This is shown in Table 4. Note that negative values of the background level c_0 or the coefficients c_2 of the first term of the asymptotic series are impossible if the profile represents an X-ray diffraction line profile.

(N,M)	(1,0)	(2,1)	(3,2)	(4,3)	(5,4)
c_0	-101	157	212	131	68
c_2	2864	-2047	-3498	-745	1899

Table 4 Values of c_0 and c_2 for a Gaussian line profile with different numbers of even terms (N) and odd terms (M) of the asymptotic series. Truncation point Δ_{tr} and inner boundary of fit range are fixed at 5.0 w and 1.5 w respectively.The level of the background (c_0) should be 100 (see Table 1) and the value of c_2 0.

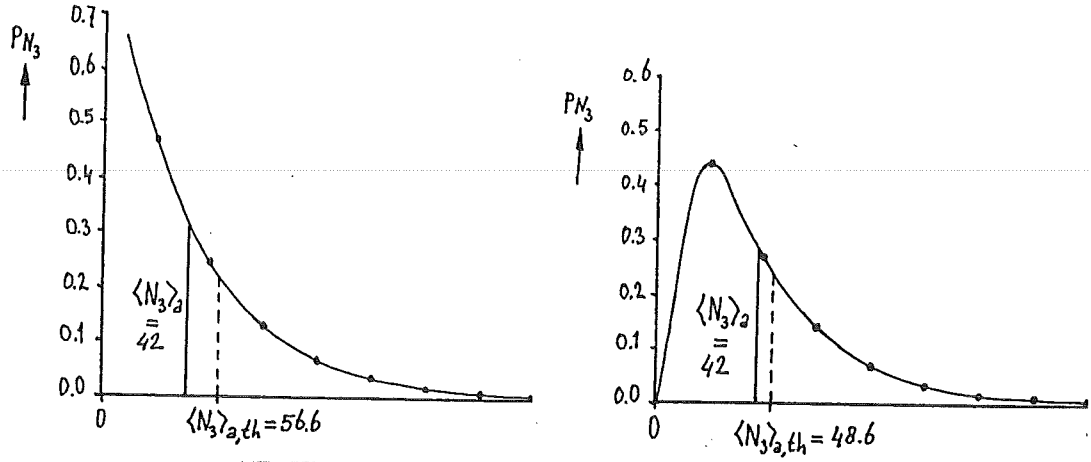

Fig. 6.2 Column length distribution of (a) Cauchy line profile (b) size-broadened line profile with $\langle N_3 \rangle_a = 42$. Dots: calculated from the sampled theoretical transform.

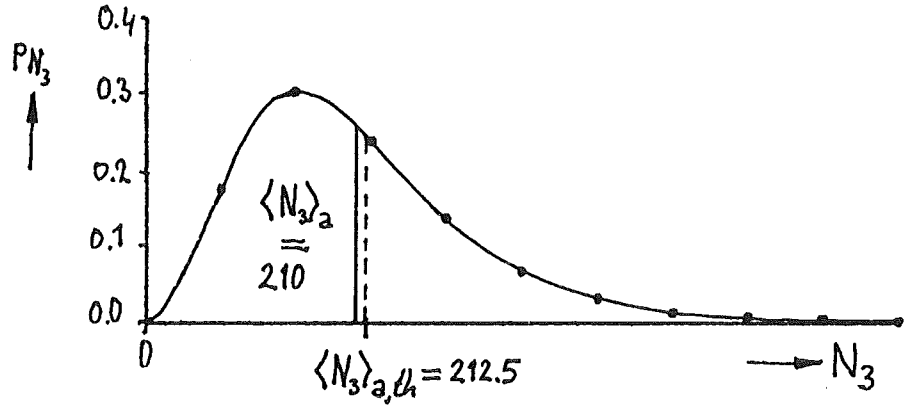

Fig. 6.3 Column length distribution of size broadened line profile with $\langle N_3 \rangle_a = 210$. Dots: calculated from the sampled theoretical Fourier transform.

Step 2 *Test with profiles for which the asymptotic series is not known, but the Fourier transform and physical characteristics are calculable.*

Two size broadened profiles with different column length distribution have been used. The column length distribution of the profile with average area-weighted column length $\langle N_3 \rangle_a = 42$ resembles the shape of the distribution of the Cauchy profile. The difference between these two distributions for small column length $\langle N_3 \rangle_a$ is negligible, because of the coarse sampling (see Fig. 6.2). It is therefore expected that the results for these two profiles show only small differences (see Table 5). The column length distribution of the size broadened profile with average column length $\langle N_3 \rangle_a = 210$ is not comparable with the distribution of the Cauchy profile. However, serious problems like those for the Gaussian profile are not expected (see Fig. 6.3).

(N,M)	(1,0)	(2,1)	(3,2)	(4,3)	(5,4)
c_2	7645	9883	10788	10815	8782
Δc_2	49	160	422	1021	2362
$\langle N_3 \rangle_{a,cor}$	51.9	47.4	46.3	46.2	48.3
$\langle N_3 \rangle_{a,th}$			48.6		

Table 5 Results of size-broadened profile ($\langle N_3 \rangle_a = 42$) with different numbers of even terms (N) and odd terms (M) of the asymptotic power series. Truncation point Δ_{tr} and inner boundary of fit range are fixed at 5.0 w and 1.5 w respectively. The theoretical value of c_2 is not known. The value of $\langle N_3 \rangle_{a,th}$ is not equal to $\langle N_3 \rangle_a = 42$, because of the coarse sampling, in Fourier space.

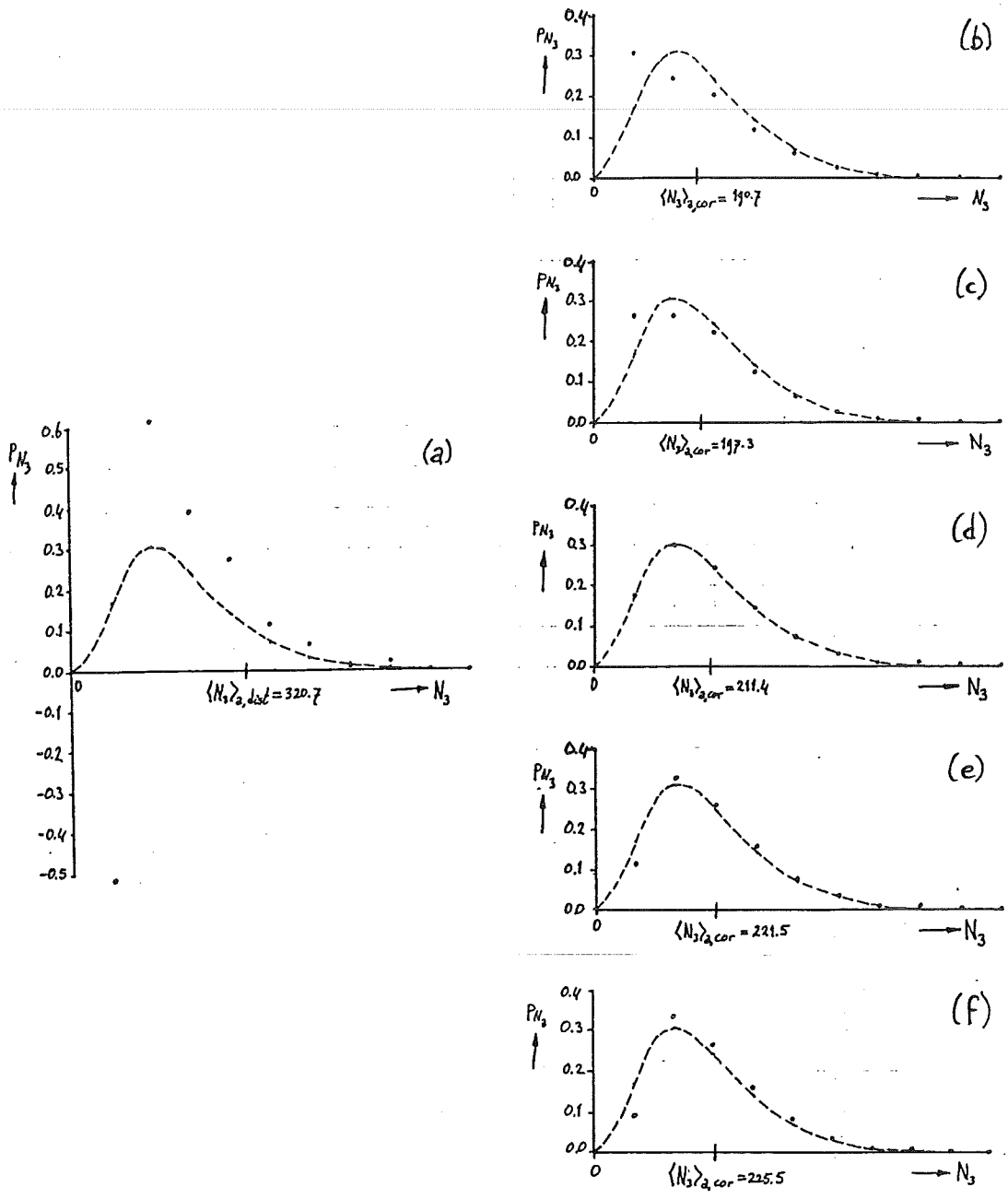

Fig. 6.4 Column length distribution of size-broadened profile ($\langle N_3 \rangle_a = 210$) (a) distorted due to truncation with average column length: $\langle N_3 \rangle_{a,dist} = 320.7$ (b)...(f) corrected with increasing numbers of even terms (N) and odd terms (M) of the asymptotic series (Cf. Table 6). Dashed line: theoretical column length distribution (Cf. Fig. 6.3).

The values obtained for the average column length shown in Table 6 are quite satisfactory when compared with the 'theoretical' value. The column length distributions calculated with the corrected DFT's show improvement on the column length distribution as calculated with the distorted DFT (see Fig. 6.4). With increasing numbers of even terms (N) and odd terms (M) of the asymptotic series the theoretical column length distribution (Cf. Fig. 6.3) is approximated more accurate. However, overcompensation occurs at (N,M) larger than (3,2).

(N,M)	(1,0)	(2,1)	(3,2)	(4,3)	(5,4)
c_2	6533	5931	4588	3645	3271
Δc_2	43	135	351	844	1945
$\langle N_3 \rangle_{a,cor}$	190.7	197.3	211.4	221.5	225.5
$\langle N_3 \rangle_{a,th}$			212.5		

Table 6 Results of size-broadened profile ($\langle N_3 \rangle_a = 210$) with different numbers of even terms (N) and odd terms (M) of the asymptotic power series. Truncation point Δ_{tr} and inner boundary of fit range are fixed at 5.0 w and 1.5 w respectively. The theoretical value of c_2 is not known. The value of $\langle N_3 \rangle_{a,th}$ is almost equal to $\langle N_3 \rangle_a = 210$, because "coarse" sampling in Fourier space has negligible effect for this column length distribution.

The effect of counting statistics on the value of the fitted coefficient c_2 is studied with 10 simulations (see Table 7). The mean value and the standard deviation are calculated. For (N,M) = (1,0) this standard deviation is comparable with Δc_2 of the fitting procedure. However, with increasing number of terms (N,M) of the asymptotic series both values increase, but Δc_2 increases faster than the standard deviation of the 10 simulations.

One can speculate that the non-orthogonality of the asymptotic series causes (i) the overcompensation reported for the column length distribution above and (ii) the difference of Δc_2 and the standard deviation of the 10 simulations. Decrease of the number of degrees of freedom probably causes the increase of the standard deviation and Δc_2.

(N,M)	(1,0)	(2,1)	(3,2)
	6540.7	6236.0	4402.3
	6442.7	6083.3	4500.6
	6503.7	5961.7	4306.6
	6459.3	6038.0	4700.0
c_2	6530.2	6115.1	4144.7
	6463.7	6088.0	4429.3
	6516.6	6151.3	4389.1
	6550.9	5953.3	4379.5
	6459.6	6034.3	4047.6
	6533.3	5931.2	4587.9
mean	6500.1	6059.2	4388.8
st.dev.	$\pm$ 40.1	$\pm$ 95.7	$\pm$192.5
Δc_2	42.8	135.7	353.3

Table 7 Effect of counting statistics on fitted coefficient c_2 of 10 simulations of the size-broadened profile ($\langle N_3 \rangle_a = 210$) with different numbers of even terms (N) and odd terms (M) of the asymptotic series. Typical values of Δc_2 (standard deviation calculated with the variance-covariance matrix of the χ^2 fitting procedure) are given in the last row.

Step 3 *Optimization of the fit.*

In chapter 4 it has already been mentioned that the fit can be optimized with respect to choice of origin, fit ranges and number of terms of the asymptotic series. A criterion is required for finding the "best" fit. The goodness-fo-fit parameters Q and R -briefly explained in Appendix K- can be used as a measure for the "average" and "systematic" misfit respectively. If both parameters exceed their threshold values, then the fit is accepted, else another origin, fit range or number of terms have to be chosen.
The use of the goodness-of-fit parameters Q ("average" misfit) and R ("systematic" misfit) is demonstrated first.
The values and interpretation of Q and R are given in Table 8 for the experiments of Table 6 and Fig. 6.4 where the number of terms of the asymptotic series has been increased. The first fit accepted (according to the goodness-of-fit parameters Q and R) is for (N,M = 3,2) which gives the best results for average column length (Cf. Table 6) and column length distribution (Cf. Fig. 6.4).

(N,M)	(1,0)	(2,1)	(3,2)	(4,3)	(5,4)
Q	<0.001 questionable	0.025 maybe acceptable	0.210 believable	0.200 believable	0.431 believable
R	0.793 systematic misfit	1.007 no syst. misfit	1.120 lucky event	1.224 lucky event	1.345 lucky event
fit acceptable ?	no	no	yes	yes	yes

Table 8 Values and interpretation (Cf. Appendix K) of the goodness-of-fit parameters Q ("average" misfit) and R (systematic misfit) belonging to the tests of Table 6 and Fig. 6.4.

However, the subsequent fits show also acceptable goodness-of-fit parameters Q and R, but the results are less satisfactory. Overcompensation occurs: the corrected average column length and column length distribution approximate the ideal values with decreasing accuracy. Therefore it seems advisable either to stop the increase of number of terms of the asymptotic series (N,M) at the first acceptable fit or to add other checks to ensure convergence.

The influence of the choice of the fit range is demonstrated in Table 9 by changing the inner boundary of the fit range using a constant value of the outer boundary Δ_{tr} = 5.0 w. Only fits with an inner boundary larger than or equal to 1.3 w are acceptable. For an inner boundary of 1.0 w an increase of number of terms to (N,M) = (4.3) gives also an acceptable fit (see Table 10).

inner boundary	1.0 w	1.1 w	1.2 w	1.3 w	1.4 w	1.5 w
Q	<0.001	0.013	0.096	0.147	0.258	0.210
R	0.858	0.990	1.103	1.185	1.199	1.199
fit acceptable?	no	no	no	yes	yes	yes
$\langle N_3 \rangle_{a,cor}$	195.5	199.3	204.1	207.4	212.3	211.4
$\langle N_3 \rangle_{a,th}$			212.5			

Table 9 Results of goodness-of-fit parameters Q and R and average column length $\langle N_3 \rangle_a$ of size-broadened profile ($\langle N_3 \rangle_a$ = 210) with different inner boundary of fit range and fixed number of terms: (N,M) = (3,2). The outer boundary of fit range is equal to $|\Delta_{tr}|$ = 5.0 w.

The influence of the choice of origin (Cf. Appendix U) is shown by a shift of the origin (see Table 11). The fit with origin shift −0.2 w is not acceptable, but increase of number of terms to (N,M) = (4.3) gives an acceptable fit as can be seen in Table 12.

(N,M)	(1,0)	(2,1)	(3,2)	(4,3)
Q	<0.001	<0.001	<0.001	0.186
R	0.263	0.502	0.858	1.172
fit acceptable?	no	no	no	yes
$\langle N_3 \rangle_{a,cor}$	199.7	184.5	195.5	209.1
$\langle N_3 \rangle_{a,th}$		212.5		

Table 10 Results of goodness-of-fit parameters Q and R and average column length $\langle N_3 \rangle_a$ of size-broadened profile ($\langle N_3 \rangle_a$ = 210) with different number of terms (N,M) of the asymptotic series and with fixed inner boundary of fit range: 1.0 w. The outer boundary of fit range is equal to $|\Delta_{tr}|$ = 5.0 w.

origin shift	−0.2 w	−0.1 w	0	0.1 w	0.2 w
Q	0.073	0.130	0.210	0.359	0.395
R	1.117	1.161	1.199	1.054	1.037
fit acceptable?	no	yes	yes	yes	yes
$\langle N_3 \rangle_{a,cor}$	206.0	208.9	211.4	215.5	212.1
$\langle N_3 \rangle_{a,th}$			212.5		

Table 11 Results of goodness-of-fit parameters Q and R and average column length $\langle N_3 \rangle_a$ of size-broadened profile ($\langle N_3 \rangle_a = 210$) with different choice of origin and with fixed number of terms (N,M) = (3,2) of the asymptotic series.

(N,M)	(1,0)	(2,1)	(3,2)	(4,3)
Q	<0.001	0.002	0.073	0.248
R	0.201	0.924	1.117	1.25
fit acceptable?	no	no	no	yes
$\langle N_3 \rangle_{a,cor}$	190.9	193.1	206.0	221.3
$\langle N_3 \rangle_{a,th}$			212.5	

Table 12 Results of goodness-of-fit parameters Q and R and average column length $\langle N_3 \rangle_a$ of size-broadened profile ($\langle N_3 \rangle_a = 210$) with different number of terms (N,M) of the asymptotic series and with fixed choice of origin (shift = - 0.2 w).

7. CONCLUSIONS

For many types of truncated signals extending from $-\infty$ to $+\infty$ it holds that:

I. For an accurate correction for truncation more than one term of the asymptotic power series $\sum_{i=2}^{\infty} \frac{c_i}{x^i}$ (where c_i are constants and x is the distance from the chosen origin of the signal) is required in general.

II. Both tails and the background can be estimated <u>simultaneously</u> using an appropriate least-squares fitting procedure.

These two main conclusions have a number of consequences. Below those that hold in particular for the field of X-ray diffraction are given.

Ia Because the asymptotic power series holds for the complete signal, it describes both tails at the same time.

Ib The asymmetry of the signal is described by the odd terms. This implies that:
- (i) the origin can in principle be chosen arbitrarily
- (ii) the limitation to signals which have a 'centre' of symmetry (in X-ray diffraction e.g. size-broadened profiles) can be dropped.

Ic The correction proposed for truncation takes into account:
- (i) wrong estimation of background level (and slope) ('horizontal' truncation)
- (ii) the intensity lost by truncation of the tails of the profile ('vertical' truncation) using the asymptotic power series

These contributions are of the same order of magnitude. The importance of the higher order (i>2) terms of the asymptotic power series becomes most clear when trying to determine the crystallite size (column length) distribution (Cf. Fig. 6.4).

Id The discrete Fourier transform values of a truncated profile can be corrected using truncation-range independent α-functions (Cf. Chapter 2).

IIa It is possible:

(i) to test the compatibility of a model (in the present case the asymptotic power series plus a straight line) with 'measured' intensities (the measured parts of the tails of the line profile plus background) using the goodness-of-fit parameters Q ('average' misfit) and R ('systematic' misfit),

(ii) to find the model parameters and their error estimates. In this investigation a χ^2 fitting procedure is applied for these purposes.

IIc The terms of the asymptotic series are non-orthogonal functions which means interdependence between the model parameters in the fitting procedure. Only the set of even terms and the set of odd terms are orthogonal, allowing separation of symmetry and asymmetry .

III. From the tests of the method performed using simulated intensity data it follows that:

IIIa The values of the goodness-of-fit parameters Q and R vary between acceptable and unacceptable values for different numbers of terms of the asymptotic power series used or for different positions of the origin. This indicates that the optimum (usually small) number of terms and the optimum origin position can be identified. However, more checks are required.

IIIb The method has been proven to be successful for the correction of truncated Cauchy line profiles and specified size broadened line profiles. However, problems are expected and noticed when the signal has a purely Gaussian shape. For such a signal the asymptotic series only virtually exists, because all coefficients are zero.

8. SUGGESTIONS FOR FURTHER RESEARCH

1. The power of the χ^2 fitting method used will increase markedly if the asymptotic power series can be written as a combination of orthogonal functions (Cf. Conclusion IIb).
2. Additional checks for the acceptability of the results of the fitting method should be developed (Cf. Conclusions IIIa).
3. The truncation correction according to method 1 of section 4.1 should be programmed and tested.
   For a component line profile the tails can be added upto the range $(l - \frac{1}{2}, l + \frac{1}{2})$, to obtain the required sampling distance in Fourier space.
   For the tails outside the range mentioned procedure 2 of section 4.5 (correction of the distorted DFT) can be used.
4. The merits of incorporating the K_α-doublet into the model to be fitted (Vermeulen, 1988–0614a, 0616ab.) should be investigated.
5. The merits of transforming from the asymptotic power series from a (h_3-l) scale to a 2θ-scale (including K_α-doublet and angle-dependent Lorentz-polarization factor) (Vermeulen 1988–0617a, 0622ab, 0902ab.) should be investigated. The importance of this translation is that <u>only then</u> counting statistics can be taken into account, because the transformation from an equidistant 2θ-scale to an equidistant (h_3-l) scale requires an interpolation mutilating the counting errors.
6. It is possible to calculate an asymptotic power series for the 'total intensity distribution' (Delhez, Keijser, Mittemeijer & Langford, 1986) by summing the contributions of all 'component line profiles' of the h_3 reciprocal lattice row (Vermeulen, 1988–0728a, 0811a, 1989–0102ab). The series obtained holds at least for size-broadening. The computer program MAGIC requires only minor changes to allow the use of 'total intensity distribution' asymptotic power series, for χ^2 fitting purposes.
7. No asymmetric profiles have been tested so far. The simulation of profiles that include strain broadening can provide asymmetric profiles.

Table G1

881220/ACV

Component line profiles and their theoretical Fourier transforms

	component line profile	theoretical Fourier transform
Cauchy	$I(h_3-\ell) = \dfrac{2\langle N_3\rangle_a^2}{1+(2\pi\langle N_3\rangle_a(h_3-\ell))^2}$	$A(t) = \langle N_3\rangle_a \, e^{-\frac{t}{\langle N_3\rangle_a}}$
Gauss	$I(h_3-\ell) = \langle N_3\rangle_a^2 \, e^{-\pi\langle N_3\rangle_a^2(h_3-\ell)^2}$	$A(t) = \langle N_3\rangle_a \, e^{-\pi\left(\frac{t}{\langle N_3\rangle_a}\right)^2}$
size broadened	$I(h_3-\ell) = \sum_{N_3=1}^{\infty} P_{N_3} \dfrac{\sin^2\pi N_3(h_3-\ell)}{\pi^2(h_3-\ell)^2}$ $P_{N_3} = \frac{1}{C} N_3^{r^*} e^{-u^* N_3 t^*}$; $C = \frac{1}{t^{*a}} u^{*\left(\frac{r^*+1}{t^*}\right)} \Gamma\left(\frac{r^*+1}{t^*}\right)$	$A(t) = \sum_{N_3=t''}^{\infty} N_3 P_{N_3} - \lvert t\rvert \sum_{N_3=t''}^{\infty} P_{N_3}$ t'' = integer : $\lvert t\rvert < t'' \leq \lvert t\rvert+1$
'Laue'	$I(h_3-\ell) = \dfrac{\sin^2\pi\langle N_3\rangle_a(h_3-\ell)}{\pi^2(h_3-\ell)^2}$	$A(t) = \langle N_3\rangle_a - \lvert t\rvert$

APPENDIX G

Brief description of the simulation of truncated signals

Four signals (line profiles) of the 'component line profile' type are available in the present version of the computer program MAGIC: Cauchy, Gauss, 'Laue' and certain size-broadened line profiles. A survey of the definitions of these line profiles and their transforms is given in table G1.

The data constituting a simulated line profile are generated by choosing a type of line profile, the truncation point Δ_{tr} and the desired sampling distance of the truncated profile. The maximum (h_3-l) range generated is fixed at (Δ_{tr}+w) and the sampling distance is fixed at 0.1 w. For all values of this array of (h_3-l) values ($-\Delta_{tr}$–w, Δ_{tr}+w) I(h_3-l) is calculated using the line profile description chosen. Then a linear background is added: I'(h_3-l) = I(h_3-l) + c_0 + c_1 (h_3-l). Finally counting statistics is simulated as follows. Each value of I'(h_3-l) is chosen as the mean value of a Poisson distribution and with the use of a random number generator one value I"(h_3-l) from this distribution is chosen: I"(h_3-l) is I'(h_3-l) including counting statistics.

In fact the program uses a "normalized" abscis and ordinate. For details about these normalizations and use of arrays in the program see Vermeulen (1989-0116a). The profile generated at ($-\Delta_{tr}$–w, Δ_{tr}+w) is truncated to ($-\Delta_{tr}$, Δ_{tr}) after the choice of origin of the truncated profile. The surplus intensities are required, because the origins can differ to investigate the effect of a shift of origin.
To be able to compare for the 'same' simulated profile (including counting statistics) the effects of various tail lengths and sampling distances, parts of the tails can be removed and other sampling distances (multiples of the original) can be chosen.

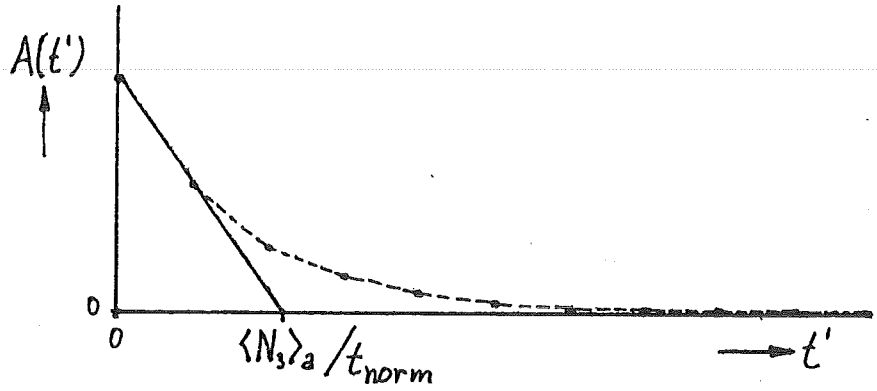

Fig. H.1 Schematic picture of the real part of the DFT A(t') versus t'.
The average area-weight column length $\langle N_3 \rangle_a$ follows from the t'-axis intercept of the tangent to the A(t') curve at t'=0.

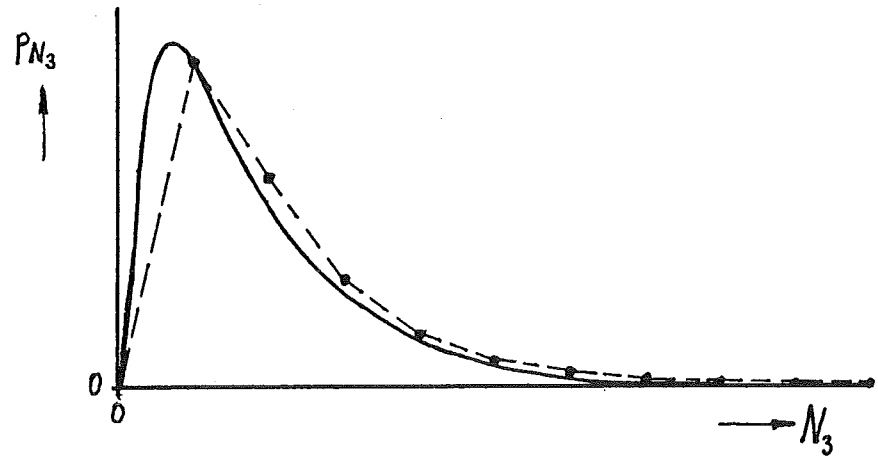

Fig. H.2 Schematic picture of column length distribution p_{N3} of profile with small average column length. Solid line: sampled at integer values of N_3 (or t). Dots and dashed line: sampled at integer values t', plotted at corresponding t values.

APPENDIX H

The calculation of $\langle N_3 \rangle_a$ and p_{N_3} for size-broadened profiles.

The average area-weighted column length $\langle N_3 \rangle_a$ and the column length distribution p_{N_3} are both calculated from the real part (cosine) of the discrete Fourier transform (DFT) values A(t'). The apparent average column length follows from (Cf. Warren, 1969) (see Fig. H.1):

$$\langle N_3 \rangle_a = \frac{A(t'=0)}{A(t'=0) - A(t'=1)} \cdot t_{norm} \qquad (H1)$$

where $t_{norm} = (2\Delta_{tr})^{-1}$ (Δ_{tr} = truncation point).
The normalization factor t_{norm} is necessary because in the program the DFT values at t' values are used instead of those at t values, with $t = t' \cdot (2\Delta_{tr})^{-1}$ (Cf. chapter 2). Eq. (H1) gives a too large value of $\langle N_3 \rangle_a$ when $t_{norm} > 1$, which is usually the case in practice.

The column length distribution is calculated with the second differences of A(t') weighted such that $\sum_{N_3=1}^{\infty} p_{N_3} = 1$ (Cf. Young, Gerdes & Wilson, 1967)

$$p(N_3 = t' . t_{norm}) = \frac{A(t'+1) - 2A(t') + A(t'-1)}{A(t'=0) - A(t'=1)} \qquad (H2)$$

The column length distribution calculated with Eq. (H2) is not determined at the integer values of N_3 or t but at the values t' with the same normalization factor as mentioned above. This will cause a loss of detail especially for line profiles with small average column length as illustrated for the case in Fig H.2.

APPENDIX K

The goodness-of-fit parameters Q and R

The goodness-of-fit parameter Q is used as a statistical estimator of the "average" misfit. Q is equal to the value of the incomplete gamma function for $\nu/2$ and $\chi^2/2$. ν is the number of degrees of freedom:

$$\nu = N - M \qquad \text{(K1)}$$

(here N = number of fit data points, M = number of estimated parameters $a_1 \ldots a_M$) and χ^2 is defined as

$$\chi^2 = \sum_{i=1}^{N} \left[\frac{y_i - f(x_i; a_1 \ldots a_M)}{\sigma_i} \right]^2 \qquad \text{(K2)}$$

(x_i, y_i, σ_i = abscis, ordinate and standard deviation of ordinate of the i-th data point respectively; $f(x_i; a_1 \ldots a_M)$ = value of the fitted function for x_i and estimated parameters $a_1 \ldots a_M$). The least squares procedure minimises the value of χ^2.

The exact description of the calculation and interpretation of Q with the incomplete gamma function is given by Press et.al., 1986. The value obtained for Q can be interpreted as follows:

if	0.0	$< Q <$ 0.001	questionable model
if	0.001	$< Q <$ 0.1	fit may be acceptable
if	0.1	$< Q <$ 0.9	fit believable
if	0.9	$< Q <$ 1.0	fit "too good".

The goodness-of-fit parameter R is used as a statistical estimator of the "systematic" misfit. The value of R -also called the successive differences scatter test- is calculated with (Thijsse, 1988 ab):

$$R = \frac{1}{2} N \, \frac{\sum_{i=i}^{N-1} (d_{i+1} - d_i)^2}{N \sum_{i=i}^{N} d_i^2 - \left(\sum_{i=1}^{N} d_i \right)^2} \qquad \text{(K3)}$$

$$\text{with } d_i = \frac{y_i - f(x_i;a_1 \ldots a_M)}{\sigma_i}. \qquad \text{(K4)}$$

The parameter R is normally destributed with a mean expected value of 1.0 and a standard deviation of:

$$\sigma_R = \sqrt{\frac{N-2}{N^2-1}} \qquad \text{(K5)}$$

The threshold value of R for accepting the "systematic" misfit is 1.0 times the standard deviation σ_R less than the mean expected value of 1.0.

The value obtained for R can be interpreted as follows:

		R	$< 1.0 - \sigma_R$	: systematic misfit
$1.0 - \sigma_R$	$<$	R	< 1.0	: no systematic misfit
1.0	$<$	R	$< 1.0 + \sigma_R$	: good result
$1.0 + \sigma_R$	$<$	R		: "lucky event".

APPENDIX Q

Calculation of the discrete Fourier transform of the lost part of the profile.

The Fourier transform of the lost part of the profile can be calculated from the intensity distribution $I_{lost}(h_3\text{-}l)$ given by Eqs. (6ab) of chapter 2:

$$F_{lost}(t) = \frac{2}{K}\int_{o}^{\infty} I_{lost}(h_3\text{-}l)\exp[-2\pi i(h_3\text{-}l)]\, d(h_3\text{-}l) \qquad (Q1)$$

With

$$F_{lost}(t) = A_{lost}(t) + iB_{lost}(t) \qquad (Q2)$$

it can easily be obtained that (Vermeulen, 1988–0525a,0705a)

$$A_{lost}(t) = \frac{2}{K}\int_{o}^{\Delta_{tr}} \sum_{k=1}^{\infty} \frac{c_{2k}}{\Delta_{tr}^{2k}} \cos 2\pi t(h_3\text{-}l)\, d(h_3\text{-}l) +$$

$$+\frac{2}{K}\int_{\Delta_{tr}}^{\infty} \sum_{k=1}^{\infty} \frac{c_{2k}}{(h_3\text{-}l)^{2k}} \cos 2\pi t(h_3\text{-}l)\, d(h_3\text{-}l) \qquad (Q3a)$$

$$B_{lost}(t) = -\frac{2}{K}\int_{o}^{\Delta_{tr}} (h_3\text{-}l) \sum_{m=1}^{\infty} \frac{c_{2m+1}}{\Delta_{tr}^{2m+2}} \sin 2\pi t(h_3\text{-}l)\, d(h_3\text{-}l) +$$

$$-\frac{2}{K}\int_{\Delta_{tr}}^{\infty} \sum_{m=1}^{\infty} \frac{c_{2m+1}}{(h_3\text{-}l)^{2m+1}} \sin 2\pi t(h_3\text{-}l)\, d(h_3\text{-}l) \qquad (Q3b)$$

Two cases of each integral of Eqs (Q3ab) will be considered: for t=0 and for t≠0. Because numerical calculations the discrete Fourier transform is required a change from the Fourier transform to the discrete Fourier transform will be performed with the substitution:

$$t = \frac{t'}{2\Delta_{tr}} \qquad (Q4)$$

where t' takes integer values only and $2\Delta_{tr}$ is the range of the discrete Fourier transformation equally taken to the truncation range.

$\underline{A_{lost}(t)\ 1^{st}\ \text{integral of Eq. (Q3a)}}$

For the case t=0 this integral reduces to

$$\frac{2}{K}\sum_{k=1}^{\infty}\frac{c_{2k}}{\Delta_{tr}^{2k}}\int_{o}^{\Delta_{tr}} d(h_3-l) = \frac{2}{K}\sum_{k=1}^{\infty}\frac{c_{2k}}{\Delta_{tr}^{2k-1}} \qquad (Q5)$$

For the case t≠0 this integral becomes:

$$\frac{2}{K}\sum_{k=1}^{\infty}\frac{c_{2k}}{\Delta_{tr}^{2k}}\int_{o}^{\Delta_{tr}} \cos 2\pi t(h_3-l)\, d(h_3-l) = \frac{2}{K}\sum_{k=1}^{\infty}\frac{c_{2k}}{\Delta_{tr}^{2k}}\frac{\sin 2\pi t\Delta_{tr}}{2\pi t} \qquad (Q6)$$

After substitution of Eq. (Q8) this becomes zero:

$$\frac{2}{K}\sum_{k=1}^{\infty}\frac{c_{2k}}{\Delta_{tr}^{2k-1}}\frac{\sin \pi t'}{\pi t'} = 0 \qquad (Q7)$$

$\underline{A_{lost}(t)\ 2^{nd}\ \text{integral of Eq. (Q3a)}}$

For the case t=0 this integral reduces to

$$\frac{2}{K}\sum_{k=1}^{\infty} c_{2k}\int_{\Delta_{tr}}^{\infty}\frac{d(h_3-l)}{(h_3-l)^{2k}} = \frac{2}{K}\sum_{k=1}^{\infty}\frac{1}{2k-1}\frac{c_{2k}}{\Delta_{tr}^{2k-1}} \qquad (Q8)$$

For the case t≠0 this integral becomes after substitution of Eq. (Q4):

$$\frac{2}{K}\sum_{k=1}^{\infty} c_{2k}\int_{\Delta_{tr}}^{\infty}\frac{\cos \pi t'\frac{(h_3-l)}{\Delta_{tr}}}{(h_3-l)^{2k}}\, d(h_3-l) \qquad (Q9)$$

The integral of Eq. (Q9) can not be solved analytically. For convenience we introduce a function $\alpha_k(t')$ as defined in appendix R which is independent of Δ_{tr}. Then Eq. (Q9) becomes:

$$\frac{2}{K} \sum_{k=1}^{\infty} \frac{c_{2k}}{\Delta_{tr}^{2k-1}} \alpha_k(t') \qquad (Q10)$$

Blost (t) 1st integral of Eq. (Q3b)

For the case t=0 this integral reduces to zero. For the case t≠0 this integral becomes :

$$-\frac{2}{K} \sum_{m=1}^{\infty} \frac{c_{2m+1}}{\Delta_{tr}^{2m+2}} \int_{o}^{\Delta_{tr}} (h_3-l) \sin 2\pi t(h_3-l)\, d(h_3-l) \qquad (Q11)$$

which can be integrated by parts yielding:

$$-\frac{2}{K} \sum_{m=1}^{\infty} \frac{c_{2m+1}}{\Delta_{tr}^{2m+2}} \frac{1}{(2\pi t)^2} [\sin 2\pi t\Delta_{tr} - 2\pi t\Delta_{tr} \cos 2\pi t\Delta_{tr}] \qquad (Q12)$$

Substitution of Eq. (Q4) gives for the discrete Fourier transform:

$$-\frac{2}{K} \frac{(-1)^{t'+1}}{(\pi t')} \sum_{m=1}^{\infty} \frac{c_{2m+1}}{\Delta_{tr}^{2m}} \qquad (Q13)$$

Blost (t) 2nd integral of Eq. (Q3b)

For the case t=0 this integral reduces also to zero. For the case t≠0 the integral becomes:

$$-\frac{2}{K} \sum_{m=1}^{\infty} c_{2m+1} \int_{\Delta_{tr}}^{\infty} \frac{\sin 2\pi t\,(h_3-l)}{(h_3-l)^{2m+1}} d(h_3-l) \qquad (Q14)$$

Integration by parts yields

$$-\frac{2}{K}\sum_{m=1}^{\infty}\frac{c_{2m+1}}{2m}\left[\frac{\sin 2\pi t\Delta_{tr}}{\Delta_{tr}^{2m}}+(2\pi t)\int_{\Delta_{tr}}^{\infty}\frac{\cos 2\pi t(h_3-l)}{(h_3-l)^{2m}}\,d(h_3-l)\right] \qquad (Q15)$$

Substitution of Eq. (Q4) gives for the discrete Fourier transform:

$$-\frac{2}{K}\sum_{m=1}^{\infty}\frac{c_{2m+1}}{2m}\left[\frac{\sin\pi t'}{\Delta_{tr}^{2m}}+\frac{\pi t'}{\Delta_{tr}}\int_{\Delta_{tr}}^{\infty}\frac{\cos\pi t'\frac{(h_3-l)}{\Delta_{tr}}}{(h_3-l)^{2m}}\,d(h_3-l)\right] \qquad (Q16)$$

The integral of Eq. (Q16) is also found in Eq. (Q9).
So Eq. (Q16) becomes:

$$-\frac{2}{K}(\pi t')\sum_{m=1}^{\infty}\frac{c_{2m+1}}{\Delta_{tr}^{2m}}\frac{\alpha_m(t')}{2m} \qquad (Q17)$$

where $\alpha_m(t')$ is defined in Appendix R.

APPENDIX R

The so-called α-functions

In this report a function $\alpha_k(t')$ is defined (Vermeulen, 1988-scriptie):

$$\alpha_k(t') = \Delta_{tr}^{2k-1} \int_{\Delta_{tr}}^{\infty} \frac{\cos \pi t' \frac{(h_3-l)}{\Delta_{tr}}}{(h_3-l)^{2k}} d(h_3-l) \qquad (R1)$$

with k and t' are positive integers, (h_3-l) and Δ_{tr} have been defined in chapter 2. The integral of Eq. (R1) cannot be solved analytically but can be reduced to a function of the Sine integral for which series expansions are known. However accurate calculation of the Sine integral is very difficult (see Appendix V for details).

Two cases of Eq. (R1) will be considered for k=1 and k > 1.

For the case k=1 Eq. (R1) reduces to

$$\alpha_1(t') = \Delta_{tr} \int_{\Delta_{tr}}^{\infty} \frac{\cos \pi t' \frac{(h_3-l)}{\Delta_{tr}}}{(h_3-l)^2} d(h_3-l) \qquad (R2)$$

Integrating by parts:

$$\alpha_1(t') = \cos \pi t' - \pi t' \int_{\Delta_{tr}}^{\infty} \frac{\sin \pi t' \frac{(h3-l)}{\Delta_{tr}}}{(h_3-l)} d(h_3-l) \qquad (R3)$$

This becomes:

$$\alpha_1(t') = (-1)^{t'} - \pi t' \left[\frac{\pi}{2} - \mathrm{Si}(\pi t') \right] \qquad (R4)$$

where

$$Si(\pi t') = \int_0^{\pi t'} \frac{\sin x}{x}\, dx \qquad (R5)$$

is the Sine integral (Abramowitz & Stegun, 1968, p.p. 231–232)
For the case k>1 a recursion formula can be deduced by integrating by parts twice in the same way as with Eq. (R2):

$$\alpha_k(t') = \frac{\cos\pi t'}{2k-1} - \frac{\pi t' \sin\pi t'}{(2k-1)(2k-2)} - \frac{\Delta_{tr}^{2k-2}(\pi t')^2}{(2k-1)(2k-2)} \int_{\Delta_{tr}}^{\infty} \frac{\cos\pi t' \frac{(h_3-l)}{\Delta_{tr}}}{(h_3-l)^{2k-2}}\, d(h_3-l) \qquad (R6)$$

Using the definition (R1) for k-1 gives:

$$\alpha_k(t') = \frac{(-1)^{t'}}{(2k-1)} - \frac{(\pi t')^2}{(2k-1)(2k-2)} \alpha_{k-1}(t') \qquad (R7)$$

From the equations (R7) and (R4) follows that $\alpha_k(t')$ is independent of Δ_{tr} and is determined completely by the values of k and t'.

APPENDIX S

Variation of truncation range

In the following the formula for the integral intensity as function of the truncation range will be deduced.
The integral intensity A_i^* or the zeroth cosine Fourier coefficient A(0) of an ideal signal I_i (h_3-l) is defined as:

$$A_i^* = \int_{-\infty}^{\infty} I_i\,(h_3\text{-}l)\; d(h_3\text{-}l) \qquad \text{(S1)}$$

The integral intensity A_i^* is often used as one of the signal characteristics in practice, but cannot be calculated with Eq. (S1) because $I_i(h_3\text{-}l)$ is not known. However, for the part of the integrated intensity of the ideal signal lost due to truncation at Δ'_{tr} we have the description (Cf. Eq. 7b):

$$\begin{aligned} A_{lost}(0) &= \int_{-\infty}^{-\Delta'_{tr}} I_i(h_3\text{-}l)\; d(h_3\text{-}l) + \int_{\Delta'_{tr}}^{\infty} I_i(h_3\text{-}l)\; d(h_3\text{-}l) \\ &= 2 \sum_{k=1}^{\infty} \frac{1}{2k-1}\, \frac{c_{2k}}{(\Delta'_{tr})^{2k-1}} \end{aligned} \qquad \text{(S2)}$$

which is a function of the truncation point Δ'_{tr}.

Further, the integrated intensity of the observed signal can be calculated with:

$$\begin{aligned} A^*_{obs}\,(\Delta'_{tr}) &= \int_{-\Delta'_{tr}}^{\Delta'_{tr}} I_{obs}\,(h_3\text{-}l)\; d(h_3\text{-}l) \\ &= A_i^* - A_{lost}(0) + \int_{-\Delta'_{tr}}^{\Delta'_{tr}} \{c_0 + c_1(h_3\text{-}l)\}\; d(h_3\text{-}l) \end{aligned} \qquad \text{(S3)}$$

-assuming the background to be linear- which is also a function of the truncation point Δ'_{tr}.
Substitution of Eq. (S2) in Eq. (S3) gives:

$$A^*_{obs}(\Delta'_{tr}) = A^*_i - 2\sum_{k=1}^{N}\frac{1}{2k-1}\frac{c_{2k}}{(\Delta'_{tr})^{2k-1}} + 2c_o\,\Delta'_{tr} \qquad (S4)$$

where c_o is the background level at the origin.
With Eq. (S4) a limited number of coefficients c_{2k} of the asymptotic series, the background level and the integrated intensity A^*_i of the ideal signal can be estimated with a least-squares fitting procedure.
Analogously it can be deduced for the n-th moment:

$$M^*_{i,n} = \int_{-\infty}^{\infty}(h_3-l)^n I_i(h_3-l)\,d(h_3-l) \qquad (S5)$$

$$M_{lost,n} = \int_{-\infty}^{-\Delta'_{tr}}(h_3-l)^n I_i(h_3-l)\,d(h_3-l) + \int_{\Delta'_{tr}}^{\infty}(h_3-l)^n I_i(h_3-l)\,d(h_3-l) \qquad (S6)$$

with (Cf. Eq. 5a):

$$I_i(h_3-l) = \sum_{k=1}^{N}\frac{c_{2k}}{(h_3-l)^{2k}} + \sum_{m=1}^{M}\frac{c_{2m+1}}{(h_3-l)^{2m+1}} \qquad (S7)$$

Then the n-th moment of the observed signal can be calculated with

$$M^*_{obs,n}(\Delta'_{tr}) = \int_{-\Delta'_{tr}}^{\Delta_{tr}}(h_3-l)^n I_{obs}(h_3-l)\,d(h_3-l)$$

$$= M^*_{i,n} - M_{lost,n} + \int_{-\Delta'_{tr}}^{\Delta'_{tr}}(h_3-l)^n\{c_o+c_1(h_3-l)\}\,d(h_3-l) \qquad (S8)$$

With the first moment a formula for the centroid can be derived. The variance is related to the second moment.
Further also the maximum intensity and integral breadth can be calculated as functions of the truncation range (Vermeulen, 1988-0221a, 0530a, 0602a, 0602b, 0602c).

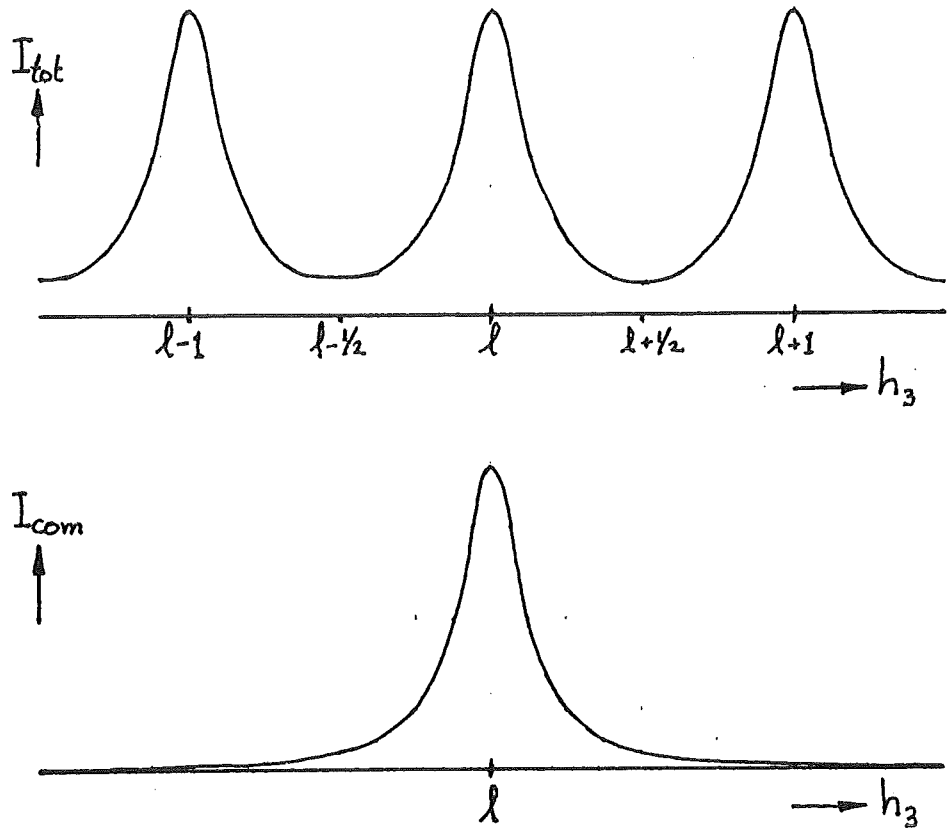

Fig. T.1 Schematic picture of the intensities I_{tot} and I_{com} -the 'total intensity distribution' and the 'component line profile' respectively- as a function of the dimensionless variable h_3 with node l (previously called reference point).

APPENDIX T

Concept of a diffraction line profile and its range

In the field of X-ray diffraction pattern analysis Delhez, Keijser, Mittemeijer & Langford (1986) regarded the 'total intensity distribution' as a series of 'component line profiles' (see also Fig. T.1):

$$I_{tot}(h_3) = \sum_{l=-\infty}^{\infty} I_{com}\,(h_3 - l) \qquad (T1)$$

For applying Fourier analysis and development of recipes for the extrapolation of the tails of the measured truncated line profile it can be considered either as part of the total intensity distribution (approach 1) or as part of a component line profile (approach 2). A summary of the two approaches is given below.

Approach 1 The total intensity distribution in reciprocal space along a [00*l*] direction is given by:

$$I_{tot}(h_3) = K \sum_{N_3=1}^{\infty} p_{N_3} \frac{\sin^2 \pi N_3 h_3}{\sin^2 \pi h_3} \qquad (T2)$$

where $h_3=(2a_3\sin\theta) / \lambda$ (a_3=lattice parameter in [00*l*] direction) θ=half of diffraction angle 2θ, λ = X-ray wavelength and p_{N_3}=fraction of columns of length -perpendicular to the 00*l* plane- N_3 (whole number) unit cells ($\sum_{N_3=1}^{\infty} p_{N_3} = 1$) and K is a proportionally constant.

Clearly the function is periodic with period from $l-1/2$ to $l+1/2$ and peaks for integer values of h_3 (see Fig. T.1).

With this approach the tails can be approximated with

$$I_{tot}(h_3) = \frac{C_{tot}}{\sin^2 \pi h_3} \qquad h_3 \rightarrow 1/2 \qquad (T3)$$

where C_{tot} is a constant.Fourier analysis implies a series development within the range $[l-1/2\ ,\ l+1/2]$.

Approach 2. The component line profile is given by

$$I_{com}(h_3\text{-}l) = K \sum_{N_3=1}^{\infty} p_{N_3} \frac{\sin^2 \pi N_3 (h_3\text{-}l)}{\pi^2 (h_3\text{-}l)^2} \qquad (T4)$$

which is not periodic in h_3, has only an absolute maximum at $h_3=l$ and entends from $-\infty$ to $+\infty$. With this approach the tails can be approximated with

$$I_{com}(h_3\text{-}l) = \frac{C_{com}}{(h_3\text{-}l)^2} \qquad (h_3\text{-}l) \rightarrow \infty \qquad (T5)$$

where C_{com} is a constant.
Fourier analysis implies a Fourier transformation within the range $[-\infty, +\infty]$.

Both approaches hold only for the case of pure size broadening. The constants C_{tot} and C_{com} have to be determined by a fitting procedure.

APPENDIX U

The influence of a change of origin

Suppose the tails of an ideal signal can be approximated with sufficient accuracy by:

$$I(h_3\text{-}l) = \frac{c_2}{(h_3\text{-}l)^2} \qquad \text{(U1)}$$

which is the first term of the asymptotic series (5a). Changing the origin from l to l' with $l' = l + \delta$ Eq. (U1) becomes

$$I(h_3\text{-}l'+\delta) = \frac{c_2}{(h_3\text{-}l'+\delta)^2} \qquad \text{(U2)}$$

The right-hand side of this equation can be expanded as an asymptotic series:

$$\frac{c_2}{(h_3\text{-}l'+\delta)^2} \equiv \sum_{k'=1}^{\infty} \frac{c_{k'+1}}{(h_3\text{-}l')^{k'+1}} \qquad \text{(U3)}$$

From Eq. (U3) it can be deduced (Vermeulen, 1988–0429a) that $c'_2=c_2$, $c'_3=-2\delta c_2$, $c'_4=3\delta^2 c_2$ etc. and

$$I(h_3\text{-}l'+\delta) = c_2 \sum_{k'=1}^{\infty} \frac{k'(-\delta)^{k'-1}}{(h_3\text{-}l)^{k'+1}} \qquad \text{(U4)}$$

To obtain an equation for $I(h_3\text{-}l)$ we substitute $l' = l + \delta$ yielding

$$I(h_3\text{-}l) = c_2 \sum_{k'=1}^{\infty} \frac{k'(-\delta)^{k'-1}}{(h_3\text{-}l-\delta)^{k'+1}} \qquad \text{(U5)}$$

Clearly the value of coefficient c_2 of the first term remains unchanged and new terms compensate the change of origin:

$$I(h_3\text{-}l) = \frac{c_2}{(h_3\text{-}l\text{-}\delta)^2} + c_2 \sum_{k'=2}^{\infty} \frac{k'(-\delta)^{k'-1}}{(h_3\text{-}l-\delta)^{k'+1}} \qquad \text{(U5)}$$

When Eq. (U5) is reduced to a finite number of terms an error $\Delta I(h_3\text{-}l)$ is introduced. The first neglected term can be used as estimate of this error. So reducing Eq. (U5) to the first term leads to an error of

$$\Delta I(h_3\text{-}l) = c_2 \frac{-2\delta}{(h_3\text{-}l\text{-}\delta)^3} \qquad \text{(U6)}$$

which is obviously systematic.

APPENDIX V

Calculation of the Sine integral

The Sine integral has to be (re)calculated numerically because existing tables and algorithms are not sufficiently accurate for the present purpose. A very considerable loss of significant decimals occurs in the recursive calculation of the α functions (see Appendix R) by applying Eq. (R7). The minimum number of significant decimals in Si(πn) and the α functions should be sufficient to obtain an accuracy better than 1% for all α values.

For the Sine integral the following convergent series expansion exists (Abramowitz & Stegun, 1968, p 232):

$$\mathrm{Si}(\pi n) = \sum_{l=0}^{\infty} \frac{(-1)^l (\pi n)^{2l+1}}{(2l+1)(2l+1)!} \qquad \text{(V1)}$$

Eq. (V1) seems to be an easily computable series. For values of the product πn less than 1 the absolute values of subsequent terms decreases and the value of Si(πn) can be obtained by direct summing until the desired accuracy is reached. However, in the present case the values of the product πn are greater than 1 -because n is a whole number- and the infinite series is still convergent, but it starts with terms that strongly increase in absolute value. The largest terms are many orders of magnitude larger than the value of Si(πn) which is about $\pi/2$. This will cause loss of significant decimals after the decimal point, because only a restricted number of digits are available for describing each term.

For large values of the product πn there also exists an asymptotic series expansion of the Sine integral (Abramowitz & Stegun, 1968, p 244):

$$\mathrm{Si}(\pi n) = \frac{\pi}{2} - \frac{(-1)^n}{\pi n}\left[1 - \frac{2!}{(\pi n)^2} + \frac{4!}{(\pi n)^4} - \ldots\right] \qquad \text{(V2)}$$

This series is divergent. However, it is suitable for numerical calculation (Erdélyi, 1956) because the series starts convergently. The best

approximation of the sum is reached by summing only the "convergent" terms i.e. until the term with the smallest value is reached.

Obviously at a certain value of the product πn the asymptotic series (V2) has to be used instead of the series (V1) because it provides a higher accuracy in numerical calculation of the Sine integral. Whether (V1) or (V2) is used in the numerical calculation, is based on an estimation of the number of significant decimals after the decimal point.

In the program the asymptotic series (V2) itself is not explicitly used for calculating the Sine integral because the asymptotic series (V2) can be combined with the equations (R4) and (R7) of Appendix R. This results in an asymptotic series for the α functions. Then the α functions can be calculated directly and the separate calculation of the Sine integral is not necessary. The asymptotic series for the α functions have the same properties as the one for the Sine integral. For more details see Vermeulen (1988–1027a).

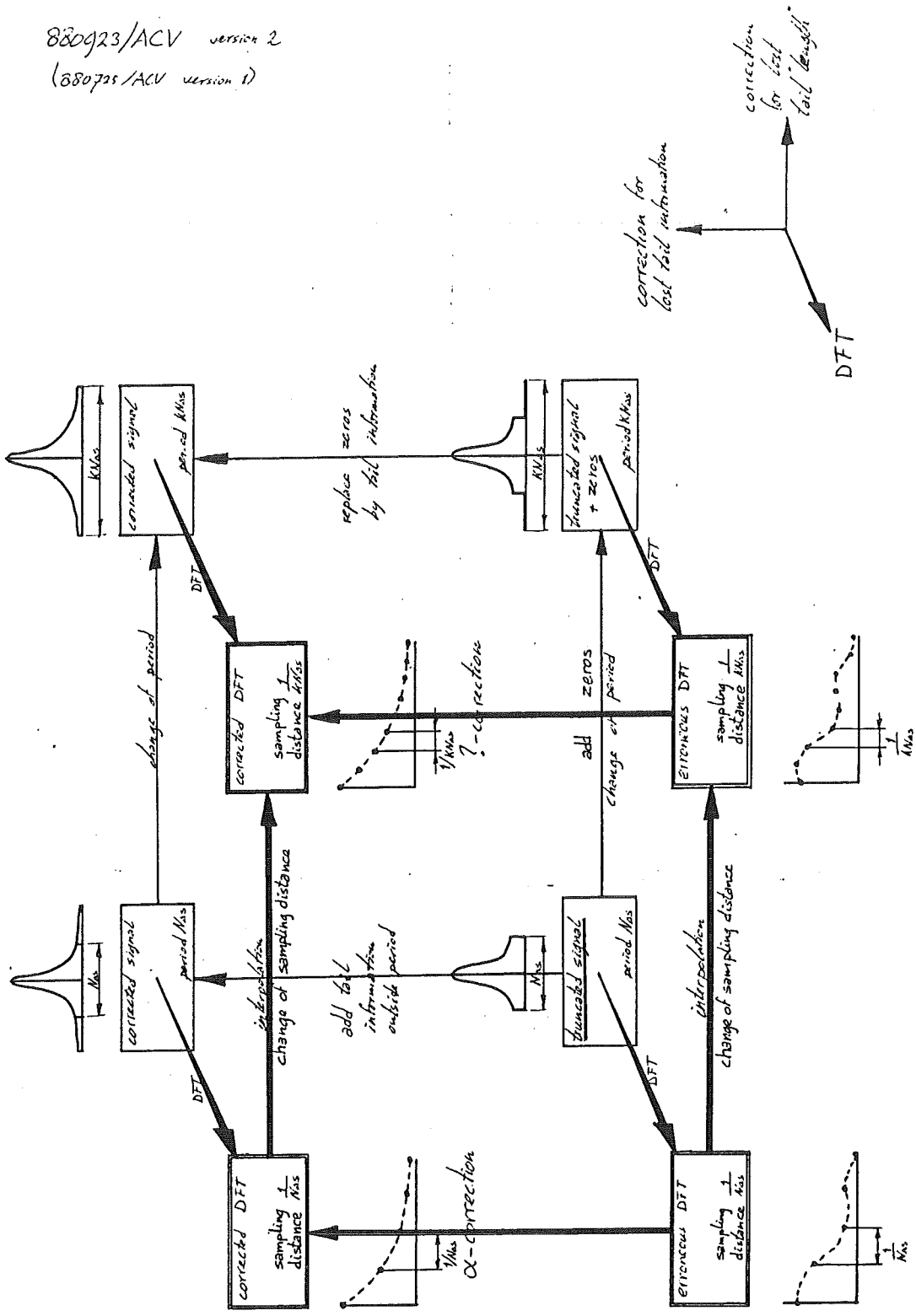
880923/ACV version 2
(880725/ACV version 1)
DFT
correction for lost tail information
correction for lost tail length
corrected signal
period KNss
KNss
replace by zeros
truncated signal + zeros
period KNss
corrected DFT
sampling distance 1/KNss
change of period
zeros
add zeros
change of period
erroneous DFT
sampling distance 1/KNss
1/KNss
?-correction
corrected signal
period Nss
Nss
interpolation
change of sampling distance
add tail information outside period
truncated signal
period Nss
interpolation
change of sampling distance
corrected DFT
sampling distance 1/Nss
erroneous DFT
sampling distance 1/Nss
1/Nss
α-correction
1/Nss

Fig. W.1

APPENDIX W

Corrections of the DFT for truncation

Many computer routines for the fast calculation of the discrete Fourier transform (see e.g.. Brigham, 1974) require a number of data equal to a power of 2. Because the measured number of data is rarely equal to such a number adding zeros is common practice to extend the number of data to the required number. However, "jumps" in the extended data set are introduced leading to an oscillatory distortion of the DFT. Adding dummy values having more plausibility instead of zeros is suggested (e.g. Bracewell, pp 374-376).
Adding data implies a change of period from $N\Delta s$ to $N'\Delta s$ (Δs = sampling distance in "measurement" space, $N'=KN$ with $K>1$) which leads to a change of the sampling distance in Fourier space from $\frac{1}{N\Delta s}$ to $\frac{1}{N'\Delta s}$. The smaller sampling distance suggest more detail, but in fact only N of the N' samples are genuine.
The influence of adding separately tail "length" and tail information on the calculated discrete Fourier transform (DFT) is considered in Fig. W1 (Cf. Vermeulen 1988-0923a). This makes it possible to show the differences between the two procedures for calculating the corrected DFT.

Procedure 1 *Adding tails*
Tail information and tail "length" are both added to the truncated signal. The corrected signal has a period of $N\Delta s$ and sampling distance of the corrected DFT is $\frac{1}{N'\Delta s}$.

Procedure 2 *Correction of the distorted DFT*
Values of the truncated signal is calculated and corrected for lost tail information using Eqs. 7abcd. Now the period of $N\Delta s$ is unchanged because no tail "length" is added and the corrected DFT keeps a sampling distance $\frac{1}{N\Delta s}$.

The change of sampling distance from $\frac{1}{N\Delta s}$ to $\frac{1}{KN\Delta s}$ can in principle also be achieved with an interpolation procedure (see Fig. W.1) For K=2 the midpoint sine interpolation formula can be used (see Bracewell, 1978 p. 427). Derivation of analogous interpolation formulae for other integer values of K must be possible, but this has not yet been investigated (Vermeulen, 1988–0725a).

REFERENCES

* Abrawowitz, M. & Stegun. I.A. (1968)
Handbook of Mathematical Functions
New York: Dover Publications

* Bracewell, R.N. (1978)
The Fourier Transform and its Applications.
Tokyo: McGraw-Hill Kogakusha

* Bringham, E.O. (1962)
The Fast Fourier Transform
Englewood Cliffs, N.J.: Prentice-Hall

* Bromwich, T.J. (1926)
Introduction to the Theory of Infinite series
London: Macmillan Co.

* Cheary, R.W. & Grimes, N.W. (1972)
The Application of Truncated Integrated Intensity to the Analysis of Broadened X-ray Diffraction Lines.
J. Appl. Cryst. (5). 57-63

* Delhez, R., Keijser, Th. H. de & Mittemeijer, E.J. (1982)
Determination of Crystallite Size and Lattice Distortions through X-ray Diffraction Line Profile Analysis.
Recipes, Methods and Comments
Fresenius Z. Anal. Chem. (312), 1–16

* Delhez, R., Keijser, Th. H. de, Mittemeijer, E.J. & Langford, J.I. (1986)
Truncation in Diffraction Pattern Analysis.
I. Concept of a Diffraction Line Profile and its Range.
J. Appl. Cryst. (19). 459-466

* Delhez, R., Keijser, Th.H. de, Mittemeijer, E.J. & Langford, J.I. (1988)
Size and Strain Parameters from Peak Profiles:
Sense and Nonsense
Aust. J. Phys. (41), 213-227

* Erdélyi, A. (1956)
Asymptotic Expansions
New York: Dover Publications

* Langford, J.I. (1982)
The Variance as a Measure of Line Broadening:
Corrections for Truncation, Curvature and Instrumental Effects.
J. Appl. Cryst. (15) 315-322

* Press, W.H. Flannery, B.P., Teukolsky, S.A. & Vetterling, W.T. (1986)
Numerical Recipes
The Art of Scientific Computing
Cambridge: Cambridge University Press

* Singleton, R.C. (1962)
An Algorithm for Computing the Mixed Radix Fast Fourier Transform.
IEEE Trans. on Audio and Electroacoustics AU 17, 93-103

* Thijsse, B.J. (1988a)
Curve Fitting
Internal paper
Delft Techn. Univ.

* Thijsse, B.J. (1988b)
Private Communication

* Vermeulen, A.C. (1988-1989)
Internal Reports, Delft Univ. Techn.

* Warren, B.E. (1969)
X-ray Diffraction
Reading, Mass.: Addison Wesley

* Wilkens, M. & Hartmann, R.J. (1963)
Zur Interpretation der Ergebnisse der Warren-Averbach-Analyse von Debye-Scherrer-Linien.
Z. Metallk., (54), 676-682

* Wilson, A.J.C. (1962)
On Variance as a Measure of Line Broadening in Diffractometry. I. General Theory and Small Particle Size.
Proc. Phys. Soc. Londen, (80), 286-294

* Young, R.A., Gerdes, R.J., & Wilson, A.J.C. (1967)
Propagation of Some Systematic Errors in X-ray Line Profile Analysis.
Acta Cryst. (22). 155-162

* Zocchi, M. (1980)
An Improved Method for the Determination of Microstructural Parameters by Diffraction-Profile Fourier Analysis
Acta Cryst. (A36), 162-170

INTERNAL REPORTS

Vermeulen, 1988–

scriptie	Algemene oplossing voor het truncatieprobleem in de Röntgendiffractie lijnprofielanalyse A.C. Vermeulen, februari 1988
0221a	De verandering van de maximale intensiteit bij de truncatie variatie methode
0429a	Invloed van foute oorsprong op machtreeks
0525a	Fout door beperken reeks bij een asymmetrisch profiel $I_V(s)$
0530a	Correctie van centroide
0602a	Bepaling maximale intensiteit en integrale breedte m.b.v. truncatievariatie
0602b	Bepaling centroide m.b.v. truncatievariatie
0602c	Halfwaardebreedte & asymmetrie
0614a	Fitten op α_2-verbrede lijnprofiel
0616a	Keuzepunten
0616b	α_2-strategieën versie 1
0617a	LP-factor meefitten
0622a	R-factor
0622b	R_S omschrijven naar θ-schaal
0705a	De discrete Fourier getransformeerd van $I_V(s)$ voor een asymmetrisch profiel versie 2
0711a	Verwijderen van ondergrond
0725a	"Sinc-interpolation"
0726a	Comparison of two methods for adding of tail information
0728a	Poster 'Total' versus 'Component'
0810a	De kristallietgrootte verdelingsfunctie van Cauchy-profiel
0811a	Het Cauchy-profiel in de 'total' & 'component' benadering
0818a	Correction of D version 3
0902a	α_2-verbrede lijnprofiel als asymptotische machtreeks op θ-schaal versie 2
0902b	Asymptotische machtreeks op θ-schaal
0923a	Poster 3-d figuur
1004a	De constante K

Vermeulen 1989–

0102a De asymptotische machtreeks in de 'total intensity distribution'-benadering volgens het sommatieprincipe

0102b Op zoek naar de asymptotische machtreeks in de 'total intensity'-benadering

0116a Description of the component line profiles and their theoretical Fourier transforms version 2