Fermi-surface reconstruction from two-dimensional angular correlation of positron annihilation radiation (2D-ACAR) data using maximum-likelihood fitting of wavelet-like functions



1998 J. Phys.: Condens. Matter 10 10507

(http://iopscience.iop.org/0953-8984/10/46/017)

J. Phys.: Condens. Matter **10** (1998) 10507–10515. Printed in the UK PII: S0953-8984(98)97431-8

# Fermi-surface reconstruction from two-dimensional angular correlation of positron annihilation radiation (2D-ACAR) data using maximum-likelihood fitting of wavelet-like functions

A G Major†, H M Fretwell†, S B Dugdale‡ and M A Alam†

† H H Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK
‡ Université de Genève, Département de Physique de la Matière Condensée, 24 quai Ernest Ansermet, CH-1211 Genève 4, Switzerland

Received 30 April 1998

**Abstract.** A novel method for reconstructing the Fermi surface from experimental two-dimensional angular correlation of positron annihilation radiation (2D-ACAR) projections is proposed. In this algorithm, the 3D electron momentum-density distribution is expanded in terms of a basis of wavelet-like functions. The parameters of the model, the wavelet coefficients, are determined by maximizing the likelihood function corresponding to the experimental data and the projections calculated from the model. In contrast to other expansions, in the case of that in terms of wavelets a relatively small number of model parameters are sufficient for representing the relevant parts of the 3D distribution, thus keeping computation times reasonably short. Unlike other reconstruction methods, this algorithm takes full account of the statistical information content of the data and therefore may help to reduce the amount of time needed for data acquisition. An additional advantage of wavelet expansion may be the possibility of retrieving the Fermi surface directly from the wavelet coefficients rather than indirectly using the reconstructed 3D distribution.

## 1. Introduction

Two-dimensional angular correlation of positron annihilation radiation (2D-ACAR) is a powerful experimental method for Fermi-surface studies of metals and alloys at a wide range of temperatures [1, 2]. 2D-ACAR experiments yield 2D projections (*Radon* transforms) of the 3D electron–positron momentum distribution, multiplied by the *momentum sampling function* (MSF) [2], also referred to as the *tent function*. In certain cases it is desirable to retrieve the 3D density distribution from the measured 2D projections.

The reconstruction of the Fermi surface from 2D-ACAR projections is an exceptionally difficult task among inverse Radon transform problems. The difficulties stem from the relatively small number of measured projections (typically less than 10). Modern methods used for this purpose, such as spherical harmonics expansion [3] or the method due to Cormack and Kontrym-Sznajd [4, 5], yield reconstructions which in many cases are in good agreement with theoretical predictions and data retrieved using other experiments. None of them, however, takes full account of the Poisson statistics of the data, and none provides a means of robustly excluding artefacts. Thus the question arises as to whether an algorithm can be constructed which takes advantage of the statistics of the data and

0953-8984/98/4610507+09$19.50  

introduces as few artefacts into the reconstruction as possible. A combination of *Haar wavelet modelling* and *maximum-likelihood fitting*, as proposed in this communication, may be an appropriate algorithm.

## 2. Haar functions and wavelets

In order to introduce the Haar functions [6], let us consider a one-dimensional experiment whose result is a set of $N$ numbers $f_1, f_2, f_3, \ldots, f_N$. These numbers can be arranged in a vector according to $\boldsymbol{f} = (f_1, f_2, f_3, \ldots, f_N)$. Information about the high-frequency and the low-frequency parts of the vector $\boldsymbol{f}$ can be obtained by calculating the differences and the sums of each pair of elements $\left[f_{2n-1}, f_{2n}\right]$, $n = 1, \ldots, N/2$ as shown in figure 1.

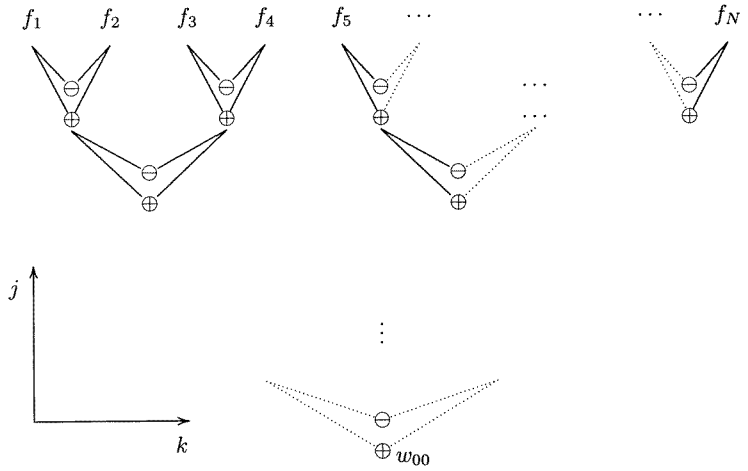


**Figure 1.** Repeated decomposition of the data in the high-frequency and the low-frequency parts.

If $N$ is a power of 2, this decomposition can be repeated with the sums while storing the differences until the row of sums contains only a single element. Let $w_{jk}$ denote the differences, where $j$ is the row in which the difference was calculated (counting from the bottom, starting at 1), and $k$ stands for the position within the row (counting from the left, starting at 0). The single sum in the last row is denoted by $w_{00}$. These $N/2 + N/4 + N/8 + \cdots + 1 = N - 1$ differences $w_{jk}$ and the single sum $w_{00}$ are referred to as the *Haar coefficients*; the transform

$$\boldsymbol{f} = (f_1, f_2, f_3, f_4, \ldots, f_N) \rightarrow \boldsymbol{w} = (w_{00}, w_{10}, w_{20}, w_{21}, \ldots)$$

is called the *Haar transform* [6]. The inverse transform can be written as a superposition of basis vectors, weighted with their corresponding Haar coefficients:

$$\boldsymbol{f} = \sum_{jk} (w_{jk} \boldsymbol{H}_{jk})$$

where the basis vectors $\boldsymbol{H}_{jk}$ are defined in real space:

$$\boldsymbol{H}_{jk} = (H_1^{(jk)}, H_2^{(jk)}, H_3^{(jk)}, H_4^{(jk)}, \ldots, H_N^{(jk)}).$$

The $N$ basis vectors $\boldsymbol{H}_{jk}$ can be interpreted as discrete functions on the interval. In the $N \rightarrow \infty$ limit, these functions converge towards the continuous *Haar functions* [6]. Except

for $H_{00}$, these functions can be derived from each other by scaling by powers of 2 and shifting along the horizontal axis. Figure 2 demonstrates the similarity of the functions and the interpretation of $j$ and $k$ as the scaling and shifting indices.

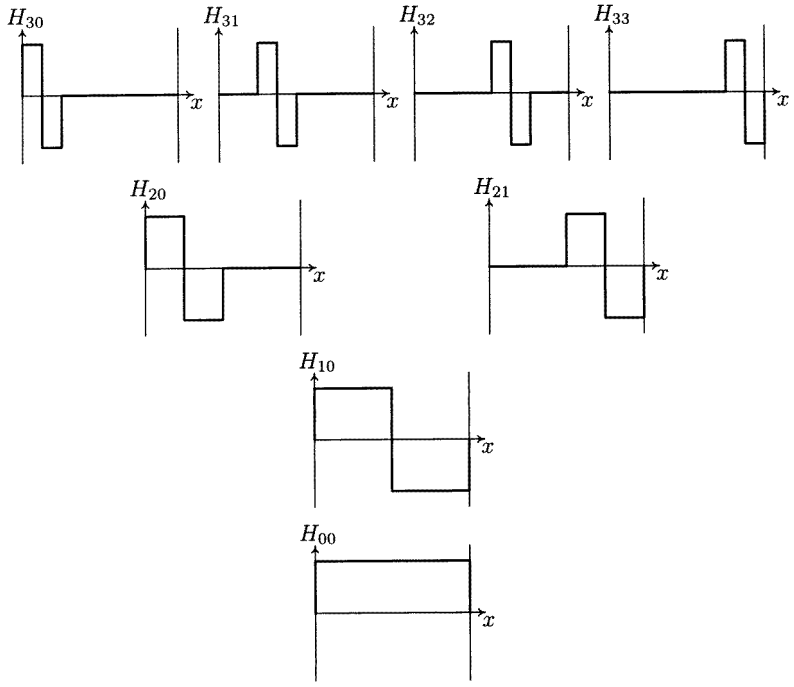


**Figure 2.** Selected Haar functions $H_{jk}(x)$. The index $j$ represents the scaling level ('size'); $k$ represents the amount of shifting ('position').

The decomposition of the original data vector $\boldsymbol{f} = (f_1, f_2, f_3, \ldots, f_N)$ can be generalized by calculating suitable linear combinations of *more than two* elements instead of sums and differences of only *two* elements in the above Haar transform scheme (see figure 1). Each choice of linear combinations yields a set of basis functions which possess the same scaling and shifting properties as the Haar function basis. These generalized Haar functions are called *wavelets* [7]; two of the most common of these are shown in figure 3. It is an interesting fact that the Haar functions (from 1910) emerge from the wavelets (1980s) as a special case.

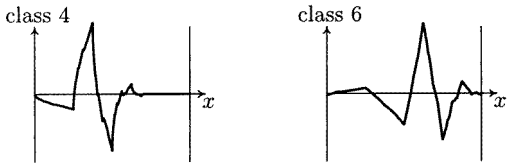


**Figure 3.** Real Daubechies class 4 (left) and class 6 (right) wavelet functions with $j = 2$, $k = 0$. The real Daubechies class 2 wavelets are identical with the Haar functions.

In many applications, wavelet analysis is superior to conventional methods, in particular Fourier analysis. Notably, wavelet functions are localized both in real space and in frequency

space, in contrast to the Fourier basis functions. As a consequence, wavelet analysis yields information both about the real-space and the frequency-space distribution of the data, while Fourier analysis loses all spatial information in favour of the frequency information. Furthermore, Fourier analysis cannot easily be applied to non-periodic data due to its periodic basis functions; this difficulty is also eliminated by the use of wavelets.

Modern applications of wavelet analysis include:

(i) solution of eigenvalue problems in quantum mechanics [8];
(ii) de-noising of experimental data [9, 10];
(iii) edge detection and pattern recognition [11];
(iv) analysis of scaling properties of fractals, DNA characterization [12] etc;
(v) image and video compression [13, 11].

Most of these applications are based on the fact that in the wavelet transforms of many non-periodic functions occurring in the real world, the majority of the wavelet coefficients are negligibly small. In the case of de-noising, these are attributed to noise; in the case of image and video compression they are assumed to be unimportant in terms of the visual impression.

Edge detection is of particular interest in 2D-ACAR analysis, where the identification of the Fermi surface is often necessary. The scaling properties of wavelets allow for efficient edge detection, as demonstrated on simulated edges in figure 4. It can be seen from the diagrams that in the highest scaling levels in the Haar transforms, a peak reflects the edge in the original data, even in the case of a deformed and blurred edge as shown on the right. It should hence be possible to detect the Fermi edge from the Haar expansion of the three-dimensional electron–positron momentum-density distribution directly, rather than from the explicitly evaluated reconstructed 3D density.

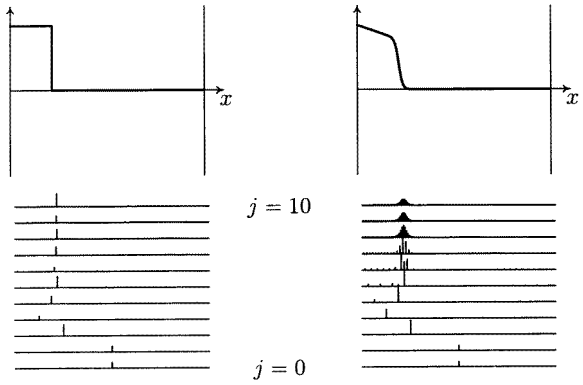


**Figure 4.** One-dimensional simulated edges (top) and their Haar transforms (bottom). Each horizontal line in the Haar transforms represents a particular scaling level $j$; the position index $k$ is identified by the position within the horizontal lines. The height of each vertical line in the Haar transforms is proportional to the magnitude of the respective Haar coefficient.

## 3. Fermi-surface reconstruction: the new approach

Most of the methods currently used for the reconstruction of the Fermi surface from 2D-ACAR data employ a *one-pass* procedure. For the particular case of Cormack's

method [4, 5], the flow diagram is shown on the left of figure 5. Initially, the 2D-ACAR projections are pre-processed: they are divided by the momentum sampling function (MSF), prior to optional feature enhancement using the maximum-entropy method (MAXENT) [14] or de-noising using wavelets [9]. With the pre-processed data, the Cormack model yields a linear set of equations. The model parameters obtained by solving the set of equations are assumed to be those that best reflect reality.

The proposed method uses a substantially different *iterative* scheme, whose flow diagram is shown on the right of figure 5. In this method, a model for the 3D electron–positron momentum density is set up. The 2D-ACAR experiment is simulated in two steps: calculation of the projections of the model and multiplication by the MSF. The theoretical projections obtained this way are compared with the experimental data, and the parameters of the model are varied such that the deviation between them is minimized. The best fit is assumed to give the best approximation to the real 3D momentum density.

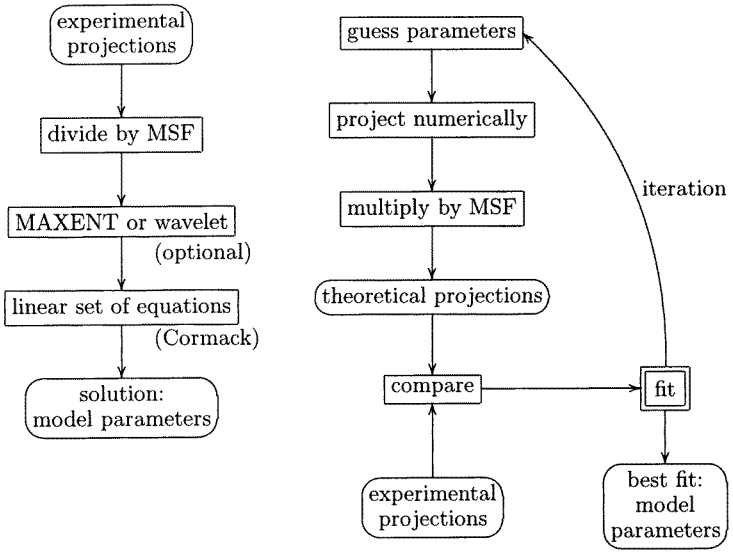


**Figure 5.** Fermi-surface reconstruction schemes. Left: Cormack's one-pass method. Right: the proposed iterative maximum-likelihood method.

The fitting approach requires the specification of

(i) the theoretical 3D model;
(ii) the method of comparing the projections.

## 4. The 3D wavelet model

We have chosen a model for the 3D electron–positron momentum-density distribution $\rho^{e^+e^-}(\boldsymbol{k})$ of the form

$$\rho^{e^+e^-}(\boldsymbol{k}) = \underbrace{a e^{-\boldsymbol{k}^2/2b^2}}_{\text{Gaussian}} + \underbrace{\sum_{\nu} p_\nu \Psi_\nu(\boldsymbol{k})}_{\text{'wavelets'}} . \tag{1}$$

The first term is an isotropic Gaussian function modelling the isotropic contribution from the annihilation of the positrons with core electrons. The second term is a sum of wavelet-like basis functions $\Psi_\nu(\boldsymbol{k})$, which are products of one-dimensional wavelet functions. These basis functions should possess the same symmetries as the reciprocal lattice of the crystal in order to minimize computational requirements. The fitting parameters of this model are the amplitude $a$ and the standard deviation $b$ of the Gaussian, and the coefficients $p_\nu$ of the wavelet-like functions.

In the case of a hexagonal lattice (as required for the investigation of GdY alloys with hcp structure), we defined $\Psi_\nu(\boldsymbol{k})$ in a cylindrical coordinate system $(r, \varphi, z)$ according to

$$\Psi_\nu(\boldsymbol{k}) = H_{j_r k_r}(r) H_{j_\varphi k_\varphi}(\varphi) H_{j_z k_z}(z) \tag{2}$$

where the functions $H_{j_r k_r}(r)$, $H_{j_\varphi k_\varphi}(\varphi)$ and $H_{j_z k_z}(z)$ are Haar functions† with their scaling indices $(j_r,\ j_\varphi,\ j_z)$ and shifting indices $(k_r,\ k_\varphi,\ k_z)$. The symmetry requirements are easily satisfied by imposing simple symmetries on the three Haar functions in equation (2).

As we have seen above, the advantage of a wavelet-like model is its capability of representing the Fermi edge with a relatively small number of elements $\Psi_\nu(\boldsymbol{k})$. In a real reconstruction, the majority of the $p_\nu$ will be negligibly small; therefore the number of simultaneous fitting parameters can be reduced significantly by an appropriate choice of elements.

## 5. Maximum-likelihood fitting

The fitting approach requires the comparison of the theoretical projections derived from the 3D 'wavelet' model with the experimental data. In our case, the number of counts in a data bin can be as low as zero; therefore the usual Gaussian approximation to the Poisson distribution would be unacceptable. Instead, a *likelihood function* $L\{m_i, d_i\}$ is defined as the probability of obtaining the given experimental data set $\{d_i\}$ from a virtual experiment whose theoretical projections are described by the expectation values $\{m_i\}$. The likelihood function for the Poisson distribution takes the form [15]

$$L\{m_i, d_i\} = \prod_i \left[ \frac{m_i^{d_i}}{d_i!} \mathrm{e}^{-m_i} \right]. \tag{3}$$

To find the best theoretical estimate, this function must be maximized with respect to the parameters of the theoretical model defining $\{m_i\}$, while the data $\{d_i\}$ are given by the experiment. Equation (3) can be transformed by defining

$$\chi^2_{\mathrm{eff}} = -2\left[\ln L\{m_i, d_i\} - \ln L\{d_i, d_i\}\right] \tag{4}$$

as the 'effective chi-square' function which must be minimized in order to obtain the best estimate. This function behaves like a Gaussian chi-square function, but is correct in the case of Poisson statistical distributions even with small count numbers in the data. Notably, the one-sigma tolerance of each parameter is given by the contour defined by $\chi^2_{\mathrm{eff}} - \chi^2_{\mathrm{eff,min}} = 1$. Equation (4) can be rewritten as

$$\chi^2_{\mathrm{eff}} = 2\sum_i \left[ m_i - d_i + d_i \ln\left(\frac{d_i}{m_i}\right) \right]. \tag{5}$$

In the fitting algorithm, we use the effective chi-square function according to equation (5) to compare the experimental data with theoretical projections. The 3D wavelet-like model,

† We could choose any wavelets here; however, Haar functions are superior to other choices, as discussed below.

as defined in equations (1) and (2), does not impose restrictions on the values of $m_i$, which can therefore become negative. In this non-physical region, $\chi^2_{\text{eff}}$ is not defined on the field of real numbers. We therefore add a penalty term to $\chi^2_{\text{eff}}$ in order to force the minimization back into the physical region.

## 6. Recent development

The proposed fitting approach offers several potential advantages over conventional methods.

(i) The statistical properties of the data are fully exploited, in contrast to the case for the majority of conventional methods, which perform well only in the limit of an extremely large number of counts. This reduces the statistics required for a good fit, hereby possibly reducing the amount of time needed for data acquisition.

(ii) The data enter the process in their original form; in particular, they are not divided by the tent function. Also, no corrections have to be applied in order to take into account any misalignment of the sample. This rotation and shifting is performed during projection of the theoretical model, eliminating the need for any interpolation.

(iii) The projection directions are not restricted to high-symmetry planes as is required for Cormack's method.

(iv) The wavelet-like function basis makes it possible to detect the Fermi surface directly from the reconstructed model parameters. In conventional methods, the explicit calculation of the 3D momentum density distribution is required.

We have implemented the proposed reconstruction algorithm as a computer program written in the C programming language. The task of minimizing the effective chi-square function is performed by the MINUIT package [16] from the CERN software library. During development, two major difficulties have arisen.

(i) Numerical integration of the 3D wavelet-like functions, which is required for calculating the Radon transforms, is very time-consuming if done by adding up values along the lines of integration. This problem was solved by restricting the choice of wavelets in the 3D model in equation (2) to Haar functions and exploiting the fact that the Haar functions can only take on three distinct values, 0, $+1$, or $-1$. This way, by calculating the boundaries of the regions in which the Haar 3D basis function is $+1$ or $-1$, integration can be done more rapidly and with high accuracy. This improvement has reduced the amount of time needed for projecting by several orders of magnitude. Without choosing Haar functions, the numerical Radon transform would be far beyond the capacities of available computers.

(ii) The number of coefficients $p_\nu$ in equation (1) can be very large (of the order of $10^4$ or more) if high resolutions are desired, making numerical minimization of the effective chi-square function impossible†. The advantage of wavelet-like functions, however, is that only a small fraction of these are required to represent the Fermi surface. Therefore, by carefully selecting the important coefficients, the number of simultaneous fitting parameters can be reduced to reasonable values. We have tried two different approaches to this non-trivial task of selecting the fitting parameters.

(a) During the runs of the minimizer, only about 20 parameters are made variable; all others are fixed. By changing the selection of these free parameters, the parameter list is scanned so that each parameter is variable in at least one run. Practice has shown that this method works reasonably well and is only moderately sensitive to the order in which the

† The MINUIT minimizer used is specified for a maximum of 50 free parameters.

free parameters are chosen. The reconstructions shown below were produced with such a strategy.

(b) The fitting procedure is started at very low resolution, so that only a small number of fitting parameters are required (up to about 100). After the first fit, the resulting parameters are compared and most of them declared insignificant due to their small magnitude. As a consequence of the scaling properties of wavelets, it is possible to tell from the results of the previous fit which of the higher-resolution parameters may be significant. Therefore, moving to higher resolutions (i.e. higher scaling levels), only a small number of new parameters must be introduced. This algorithm is more sophisticated than the first one, but also more difficult to realize.

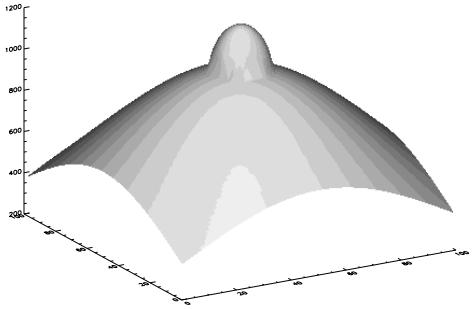


**Figure 6.** Simulated 2D-ACAR projection of an isotropic 3D electron–positron momentum-density distribution which is constant inside the Fermi sphere and vanishes outside, superposed on a Gaussian background. For simplicity, the MSF is taken to be constant over the projection.

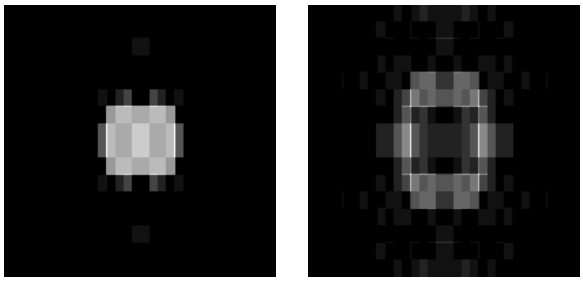

**Figure 7.** Γ–A–K cross-sections through the reconstructed 3D momentum-density distributions of the simulated isotropic material. Left: the momentum-density distribution as obtained from the fit. Right: the momentum-density distribution calculated after weighting the wavelet coefficients according to their scaling levels.

To demonstrate the capacities of the proposed method, we applied it to different data sets. We simulated the 2D-ACAR projections of an isotropic 3D electron–positron momentum-density distribution, as shown in figure 6. The model used for these test reconstructions is identical with the hexagonal model as in equations (1) and (2), but the choice of the Haar indices $j_\varphi$ and $k_\varphi$ is restricted such that the model possesses cylindrical symmetry. The results are shown in figure 7. Within the low resolution of the reconstruction ($j_{\max} = 3$ in both the $r$- and $z$-directions), the expected circle can clearly be identified. The image on the right hand side shows the possibility of detecting the Fermi edge by highlighting the

high-frequency wavelet elements.

Reconstructions from real data have not yet been successful. Our attempts to reconstruct from five projections of $Gd_{62}Y_{38}$, which were collected for use with Cormack's reconstruction method, suggest that a substantially different data acquisition strategy may be better suited for the proposed reconstruction method. As the method is not restricted to high-statistics data or given directions of projection, it may be advantageous to collect low-statistics projections in many different directions. Work on the elaboration of the proposed algorithm is in progress.

## Acknowledgments

One of the authors (AGM) wishes to express his thanks to the Studienstiftung des deutschen Volkes eV, Germany, and the Nuffield Foundation, UK, for financial support.

## References

[1] West R N, Mayers J and Walters P A 1981 *J. Phys. E: Sci. Instrum.* **14** 478
[2] West R N 1995 *Positron Spectroscopy of Solids* ed A Dupasquier and A P Mills Jr (Amsterdam: IOS) p 75
[3] Pecora L M 1988 *J. Phys.: Condens. Matter* **1** SA1
[4] Cormack A M 1963 *J. Appl. Phys.* **34** 2722
[5] Kontrym-Sznajd G 1989 *Solid State Commun.* **70** 1011
[6] Haar A 1910 *Math. Ann.* **69** 331
[7] Daubechies I 1988 *Commun. Pure Appl. Math.* **41** 909
[8] Cho K, Arias T A, Joannopoulos J D and Lam P K 1993 *Phys. Rev. Lett.* **71** 1808
[9] Major A G, Fretwell H M, Dugdale S B, Rodríguez-González A and Alam M A 1997 *J. Phys.: Condens. Matter* **9** 10 293
[10] Donoho D L 1993 *Proc. Symp. Appl. Math.* **47** 173
[11] Wu Z X and Kanamaru T 1996 *IEICE Trans. Commun.* **E79B** 1434
[12] Arnéodo A, d'Aubenton-Carafa Y, Audit B, Bacry E, Muzy J F and Thermes C 1998 *Eur. Phys. J. B* **1** 259
[13] Press W H, Teukolsky S A, Vetterling W T and Flannery B P 1992 *Numerical Recipes in C* 2nd edn (Oxford: Oxford University Press)
[14] Dugdale S B, Alam M A, Fretwell H M, Biasini M and Wilson D 1994 *J. Phys.: Condens. Matter* **6** L435
[15] Major J, Siegle A, Stammler Th, Kormann O and Major A G 1998 to be published
[16] James F 1994 *MINUIT Reference Manual* (Geneva: CERN)