The effect of regularization on the reconstruction of ACAR data

J A Weber¹, H Ceeh¹, C Hugenschmidt¹,², M Leitner¹ and P Böni¹

¹ Technische Universität München, Physik Department, Lehrstuhl E21, James-Franck-Straße, D-85748 Garching, Germany
² Technische Universität München, Forschungs-Neutronenquelle Heinz Maier-Leibnitz, Lichtenbergstraße 1, D-85748 Garching, Germany
E-mail: josef-andreas.weber@frm2.tum.de

Abstract. The Fermi surface, i.e. the two-dimensional surface separating occupied and unoccupied states in \( k \)-space, is the defining property of a metal. Full information about its shape is mandatory for identifying nesting vectors or for validating band structure calculations. With the angular correlation of positron-electron annihilation radiation (ACAR) it is easy to get projections of the Fermi surface. Nevertheless it is claimed to be inexact compared to more common methods like the determination based on quantum oscillations or angle-resolved photoemission spectroscopy. In this article we will present a method for reconstructing the Fermi surface from projections with statistically correct data treatment which is able to increase accuracy by introducing different types of regularization.

1. Introduction

When a positron is implanted into a defect-free sample, it thermalizes within the order of picoseconds. The positron will then annihilate with an electron in such a way that in the centre of mass frame two exactly collinear \( \gamma \)-quanta with an energy of 511 keV each are emitted. However, in the laboratory frame there is a Doppler shift, resulting from the longitudinal momentum of the electron-positron pair. The transverse momentum of the electron-positron pair yields an angular deviation from collinearity. By measuring both Doppler shift and angular deviation from collinearity one would be able to obtain the electron-positron momentum density (\( \rho_{ep} \)) directly. This density is continuous except for those points where bands cross the Fermi level; or in other words: It is continuous except at the Fermi surface. In general, the breaks in \( \rho_{ep} \) are the only way to identify the Fermi surface.

Due to technical limitations it is only possible to measure either the Doppler shift or the angular deviation from collinearity. With ACAR the latter is the case. This means that an ACAR measurement gives a projection of \( \rho_{ep} \). To determine \( \rho_{ep} \) it is necessary to measure a set of projections and to recalculate \( \rho_{ep} \) with the help of these. The reconstruction of a 3D density from projections is a well known problem from computed tomography (CT) and is solved by many different approaches (see e.g. [1,2]). Plenty of these methods have already been used to reconstruct \( \rho_{ep} \) from ACAR data [3,4]. We want to note, however, that there is a fundamental difference between reconstructing from x-ray images of a body and from ACAR data: As the human body has no symmetries, x-ray tomography is a 2D problem. Such is not the case with the
electronic structure of crystalline solids. Here, a projection can provide a lot more information if the full 3D symmetry is taken into account.

Since the measurement time is limited, the experimentalist has to choose whether to collect few projections with high statistics or many projections with few statistics. Different views exist on this question (compare e.g. [5], [6] and [7]). Although e.g. Kontrym-Sznajd et al. gave an rough estimation out of how many projections a 3D density can be reconstructed [8], there has been no detailed study as to which approach is optimal yet.

2. Theory

In this work we will treat the reconstruction problem by minimising \( \chi^2 \), i.e. the quadratic difference between measurement and reconstruction weighted with the statistical accuracy. Therefore some definitions are needed:

Mathematically a projection of \( \rho_{ep} \) in a direction \( \alpha \) is expressed by the so called “Radon transformation”. In the discrete case it is simply calculated by the multiplication with a projection Matrix \( P^\alpha \). Since the samples investigated with ACAR are crystalline, \( \rho_{ep} \) is highly symmetric and it is possible to reduce the complexity of \( P^\alpha \) by restricting the problem into the irreducible wedge of the reciprocal space.

In order to find a density \( \rho \) which is in agreement with the measurements, we have to minimize the function \( G(\rho) \):

\[
G(\rho) = \lambda F(\rho) + \chi^2 = \lambda F(\rho) + \sum_\alpha \sum_i \left( M^\alpha_i - \frac{\rho P^\alpha_i}{\sigma^\alpha_i} \right)^2
\]

The regularization functional \( F(\rho) \) is needed, because the problem is highly under-determined. For the reconstruction of ACAR data an entropy-like function \( F(\rho) = \sum_i \rho_i \ln \rho_i \) has been used previously [9]. This entropy function treats every point in \( \rho \) individually and tries to bring it to its lowest possible value, which minimizes the amount of additional information that is put into the reconstruction. However, we actually do have more information we can put in: \( \rho \) should have no sharp steps except for the Fermi surface breaks. A possible way to implement this is given by the sum over the gradient of \( \rho \):

\[
F(\rho) = \sum_{|\nabla \rho|_i \leq a} ((\nabla \rho)_i)^2 + \sum_{|\nabla \rho|_i > a} 2a |\nabla \rho|_i - a^2
\]

The parameter \( a \) has two purposes: It makes the regularization functional differentiable compared to a function which uses only the absolute value. Additionally \( a \) defines the strength of the regularization. This regularization is related to the algorithm proposed by Rudin et al. which uses the total variation [10].

Another very general way of putting prior knowledge of \( \rho_{ep} \) into the regularization is the histogram method. Therefore, this knowledge has to be coded into a histogram \( H \). For regularization, \( H \) is compared to the histogram of the reconstructed density \( \rho \):

\[
F(\rho) = \sum_i (H_i - \text{hist}(\rho)_i)^2
\]

In practice, instead of calculating the histogram, it is computationally less demanding to sort the point-wise values of \( \rho \). The sorted list can be interpreted as a histogram with infinitesimal resolution. Of course the problem of choosing the right histogram \( H \) for regularization still remains. This is non-trivial, but since we will work with simulated data, we can take the histogram of the original underlying density.
Table 1: Quadratic difference between original density $\rho_{ep}$ and reconstructed densities $\rho_r$ for various amounts of projections and for different regularization functions. The total number of counts in every set of projections is $2.75 \times 10^6$. For calculating the quadratic differences, the densities were normalized to the original density.

| Projections | Entropy  | Gradient | Histogram |
|-------------|---------|----------|-----------|
| 3           | $5.3 \times 10^{-7}$ | $1.3 \times 10^{-7}$ | $8.4 \times 10^{-7}$ |
| 5           | $4.0 \times 10^{-7}$ | $9.6 \times 10^{-8}$ | $4.6 \times 10^{-7}$ |
| 17          | $2.3 \times 10^{-7}$ | $7.3 \times 10^{-8}$ | $3.3 \times 10^{-7}$ |

For $\chi^2$, according to the Poisson statistic, the expected value is $\chi^2 = 1$. This can be achieved with the right choice of $\lambda$. When $\chi^2$ reaches $\chi^2 = 1$ the optimization is stopped.

To find the minimum of $G(\rho)$, a physically inspired iterative gradient approach is used: Consider $\rho$ to be the coordinates of a body that moves according to Newton’s laws of motion in the potential surface defined by $G(\rho)$, subject to friction. In every iteration step the force $F = -\nabla G(\rho)$ is calculated, which leads to the simple update formulas for $v$ and $\rho$:

$$v_i = (v_{i-1} + F \Delta t) e^{-\mu \Delta t}$$

$$\rho_i = \rho_{i-1} + v_i$$

The friction constant has to be set to $0 < \mu < \infty$ in order to ensure that the minimum will be found.

3. Results

Since the reconstruction algorithm is to be compared quantitatively, projections were created from a simulated $\rho_{ep}$ of Copper (for details see [11]). Three sets of projections with a total amount of counts of $2.75 \times 10^6$ were generated. In every set the projections were calculated by rotating the simulated density around the [110] axis in equidistant steps (22.5°, 15° and 5° respectively). The main symmetry directions at 0°, 54.7° and 90° were avoided, because these provide less information about $\rho_{ep}$ than an arbitrary projection. In a real experiment, however, at least one of this directions will be measured additionally, since it can provides valuable information about the sample orientation and quality. Three reconstructions were created per set, using a different regularization functional for each.
In figure 1 a cut from [100] to Γ to [110] through the original density is shown with the residual (original – reconstruction) for each reconstruction from five projections. Obviously the difference between original and reconstruction is small. Nevertheless it is apparent that the gradient regularization yields the reconstruction with the most efficient noise suppression and with the fewest artefacts.

A more quantitative comparison is given in table 1. There, the quadratic deviations from original density $\rho_{ep}$ and reconstructed density $\rho_r$ are listed ($d^2 = \sum_i (\rho_{ep} - \rho_r)^2$) for different methods of regularization and for different number of projections.

With an increasing number of projections $d^2$ is decreasing for every regularization method. This indicates that the measurement time is used most efficiently if many projections are recorded. But not only is the quality of the reconstruction increasing with the number of projections, the necessary iterations are decreasing. Hence, the problem of reconstruction from many projections is better determined than reconstructing from few, if the total amount of counts is the same.

On the other hand, table 1 shows the influence of regularization: While gradient and entropy regularization can increase the accuracy in every case, surprisingly the histogram regularization is not able to do so, albeit a lot of a priori information (the histogram of the original density) is used for it.

Although the entropy regularization has been the only regularization used in connection with the reconstruction from ACAR data, it is not the one which yields the best results. This is achieved with the gradient regularization in every case. The reason why the entropy regularization has been previously used so often to reconstruct ACAR data may be the simple optimization algorithm proposed by Gull and Daniell [12]. Since it takes much more time to record ACAR data than to calculate a reconstruction, low complexity or fast computation should nowadays not be the reason for choosing an algorithm. However, the optimization algorithm proposed here is nearly as simple and fast as the one of Gull and Daniell. Moreover it is a lot more flexible since the method of regularization can be changed very easily.

4. Conclusion and Outlook
In this article we tested three regularization functions on their effect of reconstructing ACAR data. A novel gradient based regularization was introduced, which was able to increase the quality of a reconstruction compared to common regularizations. Further, we have ascertained here for the first time the fact that measurement time is used most efficiently if many projections as opposed to a few projections with high statistics are recorded.

We want to point out, that the results obtained here are not only restricted to ACAR spectroscopy. In Compton scattering experiments, the measurements are also projections of the electron momentum density density [13]. Hence our results can be also applied to this kind of experiment.

Another important challenge of ACAR spectroscopy is the experimental resolution. To integrate this issue into the reconstruction algorithm is the focus of ongoing work.

Acknowledgements
This project is funded by the Deutsche Forschungsgemeinschaft (DFG) within the Transregional Collaborative Research Center TRR80 “From electronic correlations to functionality”.

References
[1] Kak A C and Slaney M 1988 Principles of Computerized Tomographic Imaging (New York: IEEE Press)
[2] Cormack A M 1963 J. Appl. Phys. 34 2722
[3] Kontrym-Sznajd G 2009 Low Temp. Phys. 35 599
[4] Weber J A, Böni P, Ceeh H, Leitner M and Hugenschmidt C 2013 J. Phys. Conf. Ser. 443 012092
[5] Haynes T D et al. 2012 New J. Phys. 14 035020
[6] Kondo H, Kubota T, Nakashima H, Murakami Y and Tanigawa S 1993 Phys. Status Solidi (b) 177 345
[7] Hamid A S et al. 2011 *Phys. Rev. B* **84**(23) 235107
[8] Kontrym-Sznajd G et al. 2002 *Phys. Rev. B* **66** 155110
[9] Pylak M, Kontrym-Sznajd G and Dobrzyński L 2011 *Appl. Phys. A* **104**(2) 587
[10] Rudin L I, Osher S and Fatemi E 1992 *Physica D: Nonlinear Phenomena* **60** 259
[11] Leitner M 2013 *Electron momentum reconstruction in ACPAR* Diploma thesis Universität Wien
[12] Gull S and Daniell G 1978 *Nature* **272** 686
[13] Utfeld C et al. 2010 *Phys. Rev. B* **81**(6) 064509