Application of MaxEnt to inverse photoemission spectroscopy

W. von der Linden, M. Donath, and V. Dose
Max-Planck-Institut für Plasmaphysik, EURATOM Association
D-85740 Garching b. München, Germany
e-mail: wvl@ibmop5.ipp-garching.mpg.de

ABSTRACT. Information about the spectral density gained by inverse photoemission spectroscopy is distorted by the Fermi distribution and the apparatus function. In many cases recovery of the desired physical quantities is hampered by an ill-posed inversion problem. It is shown, based on the spin- and temperature dependent quasiparticle spectrum of Ni, that the Maximum Entropy method yields unbiased access to the spectral density independent of model assumptions. The effective energy resolution is improved by a factor of 5, as compared to the raw experimental data, and structures below the Fermi level, which are generally lost in inverse photoemission, are recovered.

1. Introduction

In this article we will show that the Maximum Entropy Method [1] is an ideal data-analysis tool for recovering "hidden information" from experimental data without making any model-assumptions. We will address a longstanding problem in the field of itinerant magnetism. For the microscopic understanding of collective magnetism of itinerant electrons, as in transition metals, the temperature and spin dependence of the spectral density $A_{k\sigma}(\omega)$ close to the Fermi level play a vital role. In an homogeneous magnetic field the electronic spectral density consists of a single δ-function $\delta(\omega - \omega_{k\sigma})$ for given spin direction σ and momentum k. The quasiparticle energy, $\omega_{k\sigma} = \varepsilon_k + \alpha B\sigma$, depends on the free electron dispersion ε_k and the Zeeman term which splits spin-up and spin-down energies proportional to the external magnetic field. Inside a transition metal, below the Curie temperature (T_C), there exists an effective magnetic field which is proportional to the net magnetization of all electrons, $B_{\text{eff}} \propto \langle S_z \rangle$. Within the mean field approximation, B_{eff} acts like an external field and one expects $A_{k\sigma}(\omega)$ to show a pair of peaks, one for each spin direction. More elaborate approximations to the many-body problem allowing for changes in the electronic spin due to electron-electron interactions. Therefore an electron with initial spin σ experiences also states of opposite spin. One would therefore expect a "multiband structure" with temperature dependent pole strength and quasiparticle energies. With increasing temperature correlation effects are expected to lead to a mixing of spin-up and spin-down states resulting in "extraordinary" peaks. Above T_C the spin asymmetry disappears, since the rotational symmetry is restored. The multiband structure is, however, retained at and above T_C, owing to short-range ferromagnetic spin-correlations. These ideas are underpinned by, for example the fluctuating band theory [2] or approximate many-body calculations based on Hubbard-type model hamiltonians [3] and cluster-calculations [4]. However, at present there is no theory for band magnetism which is generally accepted. Even the more fundamental question, whether model hamiltonians such as the Hubbard model describe ferromagnetism at all is not settled [5]. It is therefore important to have accurate and conclusive experimental data to test the various theories.

Experimentally the situation cannot be solved rigorously either, since $A_{k\sigma}(\omega)$ cannot be measured directly. The deconvolution of the experimental IPE data is hampered by an ill-posed inversion problem. There exists an infinity of possible solutions consistent with the experimental data within the error bars.

Experimentally, one attempts to reveal the detailed spin and temperature dependence of the electronic states. A number of photoemission and inverse photoemission (IPE) studies on Fe and Ni have been performed [6, 7, 8]. While for Fe clear evidence has been found for non-collapsing band behavior at specific points in $\mathbf{k}$-space [6], the situation is more subtle for Ni and direct conclusions from the raw experimental data are not possible.

2. Formalism

To reveal such detailed features of the spectral density, as the quasiparticle energy and lifetime, particularly for states lying below the Fermi energy, which are buried under the Fermi distribution, a more subtle analysis of the experimental data is required. To this end we invoke the Maximum Entropy (MaxEnt) method which is based on Bayesian probability theory, the importance of which has been emphasized recently by P.W. Anderson [9].

The experimental IPE intensities for 100% spin-polarized electrons of spin σ are proportional to

$$I^\sigma(\omega, T, \mu) = j^\sigma \int A^\sigma(\omega') \{1 - f(\omega', T, \mu)\} g(\omega' - \omega) d\omega' \qquad . \tag{1}$$

Here $A^\sigma(\omega')$ is the required spectral density of quasiparticles with experimentally specified energy ω', spin σ and wavevector $\mathbf{k}$; j^σ represents the current density of incoming electrons of spin σ. To derive (1) standard approximations have been made, in particular ignoring matrix element- and relaxation-effects. The information about the electronic structure is contained entirely in the electronic spectral density. Dependence on temperature T and chemical potential μ enters via the Fermi distribution $f(\omega, T, \mu) = 1/(1 + \exp((\omega - \mu)/kT)$. In (1) $g(\omega' - \omega)$ stands for the apparatus function, which is a i convolution of the energy distribution of the incoming electrons and the energy window for the detected photons. The apparatus function can be estimated quite accurately from a comparison of image-potential surface states on Ni(111) measured by IPE and two-photon photoemission. We find that the apparatus function can be approximated fairly well by a Gaussian with standard deviation $d = 195$ meV [10].

For numerical purposes we evaluate the spectral density $A_n^\sigma = A^\sigma(\omega_n)$ at discrete energies ω_n with $n = 1, 2, \ldots, N_{\text{var}}$, and interpolate it linearly between these. The integrals of the piece-wise linear $A^\sigma(\omega)$ in (1) and the exact Fermi function $f(\omega, T, \mu)$ are computed numerically. Eq.1 transforms into a set of linear equations

$$I_l^\sigma = I^\sigma(\tilde{\omega}_l, T, \mu) = j^\sigma \sum_{i=1}^{N_{\text{var}}} M_{li}\, A_i^\sigma \qquad l = 1, 2, \ldots, N_{\text{eq}} \qquad , \tag{2}$$

where $\tilde{\omega}_l$ represents the (coarser) mesh on which the experimental data are available. The spin polarization of the incoming electron beam has been estimated as $p = \frac{N_\uparrow - N_\downarrow}{N_\uparrow + N_\downarrow} \approx 0.33$

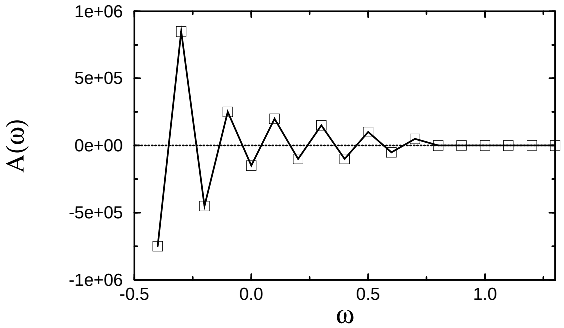

Figure 1: *Direct inversion of Eq.2 with* $N_{\rm eq} = N_{\rm var} = 20$

[8]. The incoming beam of predominantly spin-σ electrons contains, therefore, a proportion $n_{\sigma,\sigma} = (1+p)/2$ of spin-σ electrons, while the remaining proportion of electrons $n_{\sigma,-\sigma} = (1-p)/2$ has opposite spin. Therefore the measured intensity for a σ-polarized beam of incident electrons is

$$g_l^\sigma(\vec{A}) = \sum_{\sigma',i} M_{li} n_{\sigma,\sigma'} A_i^{\sigma'} \qquad . \tag{3}$$

To recover the spectral density, Eq. (3) has to be inverted. At first sight this inversion appears to be utterly ill-posed. The kernel M_{li} is almost singular due to the Fermi function, which suppresses structures below the chemical potential ($|\omega - E_F| > k_B T$) exponentially. Therefore the inverse matrix has very large eigenvalues and the experimental errors are strongly amplified. The scatter of solutions compatible to the experimental data, is therefore enormous. A direct inversion of (3) as depicted in fig. 1 (with $N_{\rm var} = N_{\rm eq}$) leads to results fluctuating between $+10^5$ to -10^5, while the real values for $A(\omega)$ are positive and of order 1. Only if the experimental data have a relative accuracy of better than 10^{-6} is direct inversion of (3) feasible. Moreover, this direct approach is restricted to $N_{\rm var} \leq N_{\rm eq}$.

We use Bayesian probability theory to determine the *posterior* probability $P(\vec{A}|\vec{g}^{\rm e},\lambda)$ for a particular solution $\vec{A}$ given the experimental data $\vec{g}^{\rm e}$ and additional experimental parameters λ, such as the scale of the error bars, the chemical potential or the width of experimental resolution:

$$P(\vec{A}|\vec{g}^{\rm e},\lambda) = P(\vec{g}^{\rm e}|\vec{A},\lambda)\,\frac{P(\vec{A}|\lambda)}{P(\vec{g}^{\rm e}|\lambda)} \qquad . \tag{4}$$

$P(\vec{g}|\vec{A},\lambda)$ is the *Likelihood* function which contains the new information provided by the experiment. In IPE experiments the data are independent and normally distributed with error σ_i. The likelihood function is therefore

$$P(\vec{g}^{\rm e}|\vec{A},\lambda) \;\;=\;\; e^{-\frac{1}{2}\chi^2} \qquad \text{with} \qquad \chi^2 = \sum_{l=1}^{N_{\rm eq}} (\frac{g_l^{\rm e} - g_l(\vec{A})}{\sigma_l})^2 \qquad .$$

Here $g_l(\vec{A})$ is the theoretically predicted result for given $\vec{A}$. The spectral density is a positive, additive distribution function, for which the appropriate entropic *prior* is invoked [11]

$$P(\vec{A}|\lambda) \quad = \quad e^{\alpha S} \quad \text{with} \quad S = \sum_i A_i - m_i - A_i \ln(\frac{A_i}{m_i}) \qquad ;$$

the information theory entropy relative to a default model m_i. We have chosen $m_i = \varepsilon$, where ε is a small quantity which serves to suppress noise in regions of insufficient information.

The MaxEnt solution for $\vec{A}$ is obtained by maximizing the posterior probability, or equivalently $\alpha S - \frac{1}{2}\chi^2$, with respect to A. The regularization parameter α is determined self consistently as elaborated by Skilling [11] upon maximizing the evidence $P(\alpha|\vec{g}^e)$ for α, given the experimental data. Other parameters, like chemical potential, width of the Gaussian resolution, and degree of polarization, can likewise be determined.

3. Discussion and Results

We have applied the MaxEnt deconvolution to temperature-dependent spin- and angle-resolved IPE data of Ni(110) for the $Z_4 \rightarrow Z_2$ transition [8]. The experimental data are taken for 20 energies per spin direction and $A_\sigma(\omega)$ is reconstructed for 80 energies. The inversion problem of equation (3) is therefore highly underdetermined. Experimental data are available for temperatures $T/T_C = 0.48, 0.64, 0.72, 0.82, 0.95$ and 1.02, covering the range from almost perfect ferromagnetic order out into the paramagnetic regime.

Before discussing the physical conclusions we will address characteristic parameters of the experiment. The statistical errors of the IPE data are known and fairly small ($\leq$ 2%). MaxEnt analysis consequently leads to a confirmation of these values. The same holds for the chemical potential, for which only slight deviations $|\Delta\mu| \lesssim 0.04$ eV from the experimentally determined values were found. A further convincing result of MaxEnt concerns the apparatus function. We allowed for more flexibility by supposing that the "Gaussian" can fall off at different rates d_l, d_r on the left and right flank of the peak. We find that the evidence is sharply peaked at a value $d_l = 183$ meV and $d_r = 195$ meV, with an uncertainty of ± 1 meV. These values are in good agreement with the estimate based on the comparison with two-photon photoemission data [10]. As the MaxEnt values for σ_l, σ_r have only very small uncertainty, this approach is very useful for determining the apparatus function whenever it is not accessible by experimental means.

Due to the incomplete spin polarization of the incoming beam one always observes two peaks in the experimental raw data. The polarization of the incoming beam had been set experimentally to make the extraordinary peak vanish for the $T/T_C = 0.48$ data [8]. Using this experimentally determined value $p \approx 33\%$ for all temperatures we find almost negligible and temperature independent "extraordinary" peaks. As MaxEnt is not a linear method it is a good idea to use it on the full experimental information in the form of Eq. 3 using p as adjustable parameter. Since the polarization of the incoming beam is independent of the sample temperature, the same polarization p is used for all temperatures and the combined evidence $\prod_i P(p, T_i)$ has to be maximized simultaneously. The maximum evidence is obtained for $p = 0.32$, which is in good agreement with the experimentally determined value

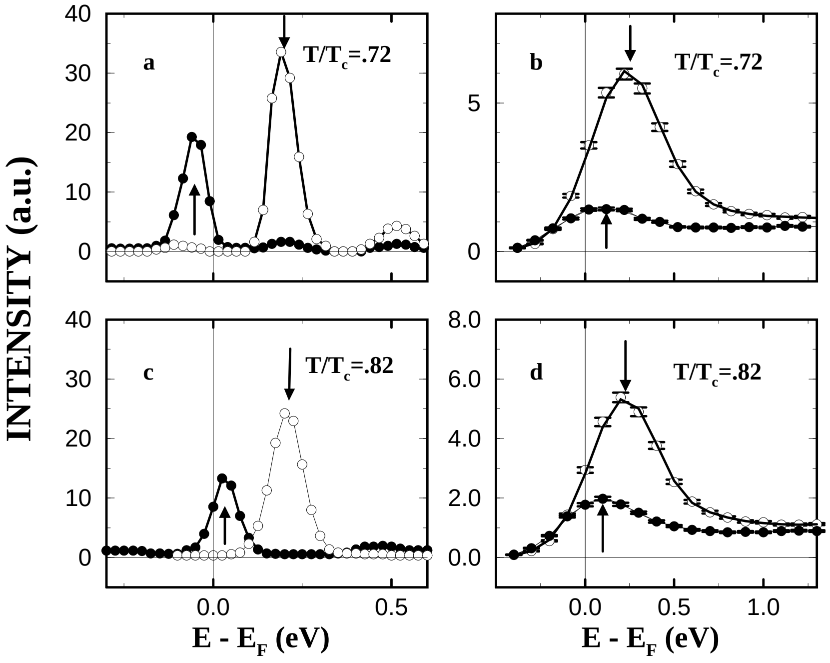

Figure 2: *Spin-dependent quasiparticle spectral density (a,c) and experimental IPE data (b,d) of the $Z_4 \to Z_2$-transition in Ni for two temperatures $T/T_C = 0.72$ (a,b) and 0.82 (c,d).*

of $p = 0.33 \pm 0.03$. At this value of p extraordinary peaks disappear for all temperatures.

Typical results obtained by the MaxEnt deconvolution are given in fig.2a,c for $T/T_C = 0.72$ and 0.82. For comparison we also depict the experimental data. In the experimental data (fig.2b,c), both spin-up and spin-down features appear above E_F [8]. The reconstructed spectral densities, however, reveal the spin-up peak clearly below (above) E_F for $T/T_C = 0.72\,(0.82)$ with a line-width of about 80 meV independent of temperature. The resolution of IPE+MaxEnt is better then 40 meV, which is an improvement by at least a factor of 5 over the raw experimental resolution. The explanation is that the experimental resolution is due to a convolution with a smooth function which can be characterized extremely accurately by a few parameters, independent of temperature. There is no significant indication of "extraordinary" peaks at all temperatures. To quantify this statement, we have determined the posterior probability of a two-peak structure, where we have mixed in a proportion q of the minority peak to the majority structure. It appears that the posterior probability falls like $P(q)/P(0) \approx e^{-aq^2}$, were the constant a depends on temperature. Remarkably, in the $T/T_C = 0.72$ data of fig.2 the posterior probability falls to $1/e$ already for $q = 0.009$ even though the results are still within the experimental error bars. This observation demonstrates emphatically that extraordinary peaks can be ruled out. With increasing temperature a decreases slightly as the peaks approach each other. For $T/T_C = 0.95$, the posterior probability drops below $1/e$ at $q = 0.02$. The slight structures visible in fig.2 for

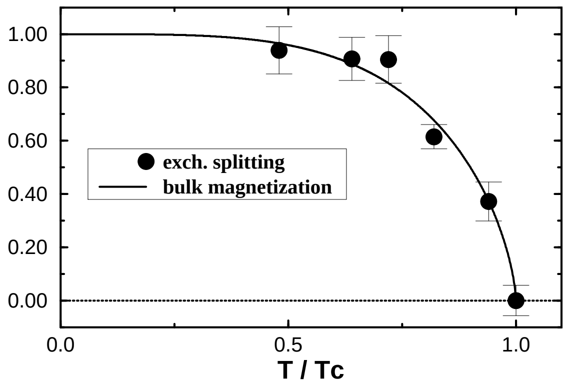

Figure 3: *Exchange splitting $\Delta E_{\rm ex}(T)/\Delta E_{\rm ex}(0)$ of the Z_2-band in Ni as a function of temperature (full circles). Errorbars are obtained selfconsistently from MaxEnt. Extrapolation yields $\Delta E_{\rm ex}(0) \approx 0.28eV$. The solid line is the experimental bulk magnetization $M(T)/M(0)$ of Ni rescaled to fit the $\Delta E_{\rm ex}$ data [12].*

$T/T_C = 0.72$ are attributed to noise and are completely absent for $T/T_C \geq 0.82$. The minor structure at ≈ 0.5 eV stems from irregularities in the experimental data a few eV above E_F and has no physical relevance.

It is instructive to compare the experimental data with those obtained by using the MaxEnt result for $A(\omega)$ in Eq. 3 (solid lines through the data points in fig. 2(b),(d)). The agreement is perfect, but for ill-posed inversion problems this is not surprising. It is likewise a completely useless test for theories to compare the theoretical and experimental values of g_l. The different heights of the peaks above and below the chemical potential should not be taken too seriously for the following reason. With a uniform model, MaxEnt reduces structures where the signal-to-noise ratio is poor, which is the case below μ due to the exponential decay of the Fermi function. The same argument leads also to a slight shift of structures in the direction, in which the kernel of the transformation increases. In the present case we therefore expect that structures below μ are actually somewhat lower in energies. This effect is, however, included in the error bars given by MaxEnt. The temperature dependence of exchange splitting $\Delta E_{\rm ex}(T)/\Delta E_{\rm ex}(0)$ for the Z_2 band in nickel is given in fig.3. The zero temperature value is estimated by extrapolation as $\Delta E_{\rm ex}(0) \approx 0.28eV$. The data follow nicely the rescaled experimental bulk magnetization curve [12] which yields strong support for a Stoner-like band behavior. The extrapolated ground-state exchange splitting of the magnetic Z_2-band is 0.28 ± 0.05 eV. Similar values, ranging from 0.17 to 0.33 eV, have been reported for occupied d-bands in Ni [13].

In conclusion, we have shown that the maximum entropy method gives spin-dependent quasiparticle spectral densities from IPE data. In the present case the resolution is improved by a factor of 5 and structures are recovered below E_F which are generally lost in inverse photoemission. This is important for the study of electronic structures in general and high temperature superconductors in particular, where the detailed behavior of quasiparticle energies and lifetimes is important for theoretical understanding.

We found that the quasiparticle spectral density of Ni consists of only one peak per spin direction for all temperatures. The exchange splitting $\Delta E_{\rm ex}$ decreases with increasing temperature and vanishes at T_C. Hence, it appears that the influence of transverse spin fluctuations is negligible for the electronic bands in Ni in the energy regime under consideration. The Maximum Entropy concept is clearly very useful to deconvolve experimental data, and can be applied immediately to other types of spectroscopy.

References

[1] "Maximum Entropy in Action", ed. B. Buck and V.A. Macaulay, *Oxford Science Publications, Oxford*, 1990.

[2] V. Korenman and R. E. Prange, *Phys. Rev. Lett.*, **53**, 186, 1984. V. Korenman, "Metallic Magnetism", ed. H. Capellmann, p. 109 *Springer, Berlin*, 1987; D.M. Edwards, *J. Magn. Magn. Mat.*, 45, 151, 1984.

[3] W. Borgiel, W. Nolting, and M. Donath, *Solid State Commun.*, 72, 825, 1989; J. Braun, G. Borstel, and W. Nolting, *Phys. Rev.*, B46, 3510, 1992.

[4] H. Gollisch and R. Feder, *Solid State Commun.*, 76, 237, 1990.

[5] W. von der Linden and D.M. Edwards, *J.Phys.* C3, 4917, 1991.

[6] E. Kisker, K. Schröder, M. Campagna, and W. Gudat, *Phys. Rev. Lett.*, 52, 2285, 1984; J. Kirschner, M. Glöbl, V. Dose, and H. Scheidt, *Phys. Rev. Lett.*, 53, 612, 1984.

[7] D. E. Eastman, F. J. Himpsel, and J. A. Knapp, *Phys. Rev. Lett.*, 40, 1514, 1978; C. J. Maetz, U. Gerhardt, E. Dietz, A. Ziegler, and R. J. Jelitto, *Phys. Rev. Lett.*, 48, 1686, 1982; H. Hopster, R. Raue, G. Güntherodt, E. Kisker, R. Clauberg, and M. Campagna, *Phys. Rev. Lett.*,51, 829, 1983; K.-P. Kämper, W. Schmitt, and G. Güntherodt, *Phys. Rev.*, B42, 10696, 1990.

[8] M. Donath and V. Dose, *Europhys. Lett.*, 9, 821, 1989.

[9] P. W. Anderson, *Physics Today*, 45, 9, 1992.

[10] W. von der Linden, M. Donath, and V. Dose, *Phys. Rev. Lett.*, 71, 899, 1993.

[11] J. Skilling, "Maximum Entropy and Bayesian Methods", ed. P. F. Fougère *Kluwer Academic Publishers*, 1990.

[12] P. Weiss and R. Forrer, *Ann. Phys.*, 5, 153, 1926.

[13] F.J. Himpsel, J.A. Knapp, and D.E. Eastman, *Phys.Rev.*, B19, 2919, 1979; P.Heimann, F.J.Himpsel, and D.E. Eastman, *Solid State Commun.*, 39, 219, 1981.