APPLICATIONS OF MAXIMUM ENTROPY TO CONDENSED MATTER PHYSICS

D. A. Drabold, A. E. Carlsson, and P. A. Fedders
Department of Physics, Washington University,
St. Louis, Missouri 63130, U.S.A.

ABSTRACT. We describe recent applications of maximum entropy to matter in condensed phases. Applications to spin systems, electronic structure, and calculating interatomic potentials are included.

I. INTRODUCTION

Maximum entropy methods have recently been applied to several kinds of problems in condensed matter physics. In broad outline these applications have fallen into either of two categories: moment problems (useful in spin systems, electronic structure calculations and densities of states for lattice vibrations), and a sophisticated application of maxent to the calculation of interatomic potentials in metals. These approaches have had considerable success, and our expectation is that extensions of these methods, and entirely new applications will be developed in the future. In solid state physics workers sometimes resort to approximations that have no *a priori* justification -- an example we will mention is the use of *ad hoc* functional forms to invert moment problems. Beside the information theoretic advantage of using maxent, another point in favor of this procedure is that it provides concrete functional forms to manipulate and base other approximations on: this can be compared to some large-scale computer calculations. It will be the goal of this paper to familiarize maxent practitioners with recent work in condensed matter and to relate a few rather generally encountered properties that might occur in other applications. We will organize this paper as follows: Section II will discuss the maxent solution of the classical moment problem and physical applications. Section III will include a brief discussion of methods for obtaining interatomic potentials via maxent.

II. MOMENT PROBLEMS IN CONDENSED MATTER

The classical moment problem may be stated as follows: Given the first N power moments of a non-negative function $\rho(x)$ on some interval $a \le x \le b$:

$$\mu_n = \int_a^b dx\, x^n \rho(x) \quad , \quad n = 0,1,2,...,N \quad , \tag{1}$$

J. Skilling (ed.), Maximum Entropy and Bayesian Methods, 137–142.

develop an approximation for ρ based on the information contained in the moments μ_n. For finite N, it is clear that the solution to the moment problem is not unique: many functional forms can be invented which correctly reproduce the known moments, but which may differ (sometimes radically) in the unknown higer moments. This lack of a unique solution leads us to consider what the "optimal" functional form might be. The answer is provided by the method of maximum entropy.[1] Following Mead and Papanicolaou, we may construct an entropy

$$S = -\int_a^b dx\, \rho(x)[\log \rho(x) - 1] \quad , \tag{2}$$

to be maximized subject to the constraint that ρ should have the required first N moments. Using the usual procedure of introducing an auxiliary functional with undetermined Lagrange multipliers to include the constraints, and functionally differentiating with respect to ρ, we easily arrive at the maxent solution of the problem:

$$\rho(x) = \frac{1}{Z} \exp\left\{-\sum_{i=1}^{N} \lambda_i x^i\right\} \quad , \tag{3}$$

and Z, λ_i are determined by requiring that ρ satisfy the moment constraints. The determination of the λ_i is difficult, owing to the nonlinearity of the maxent ρ and the many quadratures involved in an iterative procedure. A Newton minimization procedure was given in Ref. 1. Bretthorst[2] has developed a more robust algorithm, and Drabold[3] has found some sum rules that speed the original Mead and Papanicolaou code up by more than a factor of two. Despite all of this, it is not hard to find examples for which the maxent code fails. Not surprisingly, this tends to happen for functions which have explicit singularities, or those with discontinous derivatives. Such functions are sometimes of physical interest.

Before discussing examples of moment problems in condensed matter, we note that these moment problems are a recurring theme of the subject. Solid state theorists tend to work with moment formulations of problems because they offer an alternative to the task of diagonalizing large matrices. For a spin ½ problem for example, the dimensionality of the Hamiltonian matrix which contains all dynamical information is 2^K where K is the number of spins -- typically order 10^{23} for a macroscopic system! In contrast, calculation of the low order moments is usually fairly straightforward.[4] Also, the moments tend to contain information about the local environment of a particular site; information one is usually interested in. The complete set of eigenvalues and eigenvectors associated with the Hamiltonian contains vastly more information, most of which is irrelevant to an investigators specific interests. It is also worth noting in passing that the reason why one can readily extract the moments is related to the fact that the trace of a quantum mechanical operator is independent of the choice of basis.[5] In each of the examples we discuss, this is the key to obtaining the moments of physically relevant functions. In fact, with the appearance of reasonably reliable code for solving the maxent moment problem, we may think of

maxent as providing us with an alternate numerical method for diagonalizing Hamiltonian matrices: After all, we can (in principle) *always* calculate traces of powers of such matrices. These traces are to within a normalization exactly the power moments of the density of states for the Hamiltonian matrix. So for cases where symmetries or other considerations allow easy calculation of powers of the Hamiltonian, maxent should be considered as a means of obtaining the density of states. One other potential application of maxent moment methods to solid state physics is the improvement of convergence of certain expansions. For example, a high-temperature expansion takes the form of an infinite series in traces of powers of the Hamiltonian. Information theoretic extrapolation for higher order terms in the series may provide a useful means of extracting physically meaningful results for lower temperatures.

The first physical application of the moment problem we discuss is the calculation of response function $G(\omega)$ for spin systems.[6] This function has the physical interpretation of a spectral density: it may be thought of as indicating the "density of excitation" per unit frequency range. It is clear that such a function must depend upon the detailed dynamics of the spins, which is naturally a many-body quantum mechanical problem. Several kinds of spin-spin interactions have been examined: Mead and Papanicolaou[1] applied maxent to the one-dimensional XY model and Heisenberg exchange. Impressive agreement with the exact solution of the XY model was obtained. Because maxent was well converged in the Heisenberg case (in the sense that the answer did not change appreciably with additional moments), these authors reasonably concluded that they had an essentially exact solution for the spin dynamics of the Heisenberg system. This result is significant, because there is no exact solution known.

For a calculation which may be directly compared to experiment, we turn to the case of a magnetic dipolar coupling between spins ½. Here, careful experiments have been done on CaF_2 where the fluorine nuclear spins are arranged in a simple cubic lattice. Some complicated calculations[7] have produced eight exact moments for the physically measurable "lineshape" function $G(\omega)$ for this system. In Ref. 8 maxent was applied to the theoretical moments and found to agree with experiment to within ~2%. In general the function $G(\omega)$ is complex-valued, the real part representing the NMR absorption spectrum, which is clearly positive definite. It has been found convenient in other calculations [9] to introduce a function related to G, the self energy $\Sigma(\omega)$ which satisfies the equation $G(\omega) = i/\{[\omega - \Pi(\omega) + i\Gamma(\omega)]\}$, where Π and $-\Gamma$ are the real and imaginary parts of Σ respectively. It can be shown that Γ is of one sign, and therefore another candidate for the application of maxent. Power moments of Γ are readily related to the known theoretical moments of $G(\omega)$. Although the direct use of maxent on the function G was very satisfactory, optimal agreement was obtained by fitting Γ as an intermediate step. The reason for this appears to be that Γ has less structure than G and is therefore easier to apply maxent to. The utility of the auxiliary function Σ is related to the function-theoretic properties of G and Σ on the complex plane. This point is discussed further in Ref. 8. Another interesting feature of the work on the dipolar lattice was the appearance of an oscillating pattern

of convergence. As others have shown, it is quite possible to find sets of moments for which the maxent procedure does not converge. By this we do not refer to a numerical difficulty, but to an intrinsic limitation of the method for certain sets of input moments. For the dipolar case it was observed that for $N = 4k + 2$ (N the number of input moments, k an integer), no maxent solution existed on the infinite interval. This result is connected with the expected large–ω behavior of the exact G, which on physical grounds is expected to decay like $e^{-a|\omega|}$, for some a. On the other hand, the maxent fitting function behaves like $e^{-\lambda_N \omega^N}$ thus to reproduce the correct behavior, there must be considerable cancellation for large ω, implying that the Lagrange multipliers should not be of one sign. Cases for which $\lambda_N < 0$ can lead to situations for which there is no maxent solution. Numerically this non-convergence is manifested by a dependence of the λ_i on the cutoffs for the numerical integrals (the actual range of integration for G and Γ is $\omega \in (-\infty, \infty)$). It was also found that self-energy and lineshape fits were complimentary in the sense that when the lineshape (self energy) produced a non-converged calculation, the self-energy (lineshape) function converged. So in some cases, one may be forced to use an auxiliary function to obtain a converged maxent fit.

We have also applied maxent to a more complicated version of the previous problem, the case of a nonmagnetic host with spins −½ randomly diluted throughout a crystal.[10] A particular realization of such a system is ordinary diamond, due to the existence of two isotopic species of carbon: spin 0 (magnetically inert) and spin ½. For high concentrations of magnetic particles it was found that maxent and configuration averaged moments produced good line shapes. For low concentrations of spins, we used maxent as an aid in inferring to what extent spin wavefunctions were localized (in the terminology of magnetic resonance this characterized the dipolar broadening as inhomogeneous or homogeneous).

We have also recently applied maxent to the problem of obtaining theoretical estimates of relaxation times in solid molecular hydrogen. We have observed reasonable agreement between theory and experiment.[11]

Maxent has been used to obtain densities of states in binary random alloys. Here, there have been a wide range of methods applied, from exact diagaonalization of large matrices to recursion methods. For a particular model calculation[12] it was found that maxent offered a real alternative to continued fraction and coherent potential approximation (CPA) methods. While it is certainly true that the CPA method produces very satisfactory results in a wide range of regimes, it is limited in some contexts by mean-field like assumptions underlying its derivation. The formulation of recursion and maxent moment methods do not suffer from this weakness. Maxent also has an advantage over recursion; in the usual implementation of recursion the electronic Green's function takes the form of a continued fraction which must be terminated in some way. This is unfortunately more an art than a science. While an experienced practitioner of recursion would correctly argue that a particular choice of truncation schemes incorporates knowledge of the physics of the problem, we point out that such information could also be included as an additional

constraint on maxent, without the introduction of bias. As in the spin problem it was useful to reconstruct functions related to the electronic Green's function rather than the Green's function directly. This was again because the auxiliary functions were better behaved. In the alloy problem it was best to use a function which bears the same relation to the self energy that the self energy did to the Green's function in the spin problem. This procedure produced the best agreement with a CPA calculation. We should also point out that maxent and recursion are complementary to some extent, as recursion provides the most efficient means of calculating the moments needed for the maxent procedure. We also tried a specific example of one vacancy in crystalline Si, and found maxent to be superior to continued fractions.[13]

Brown and Carlsson[14] provided the first application of maxent to structural energy calculations in the presence of defects such as vacancies. Very recently a comprehensive study of methods for calculating bond energies in a tight binding model has appeared.[15] It was found that maxent was a useful means for computing structural energies in the presence of defects, better than continued fractions with a square root terminator, but only roughly equal in accuracy to a gaussian quadrature[17] approach which was computationally easier for more than six recursion levels. Glanville, *et al.*[15] observed that there were computational difficulties with maxent due to an extreme sensitivity of the maxent fit to the values of the Lagrange multipliers conjugate to the moments. A possible remedy for this difficulty is to solve the moment problem on a different basis. The origin of the trouble lies in the nearly singular nature of the covariance (Hessian) matrix -- this is a consequence of the increasing degree of correlation between higher moments. A possible solution is to take N linearly independent combinations of the moment constraints, solve the moment problem on the new constraints, and transform back. Bretthorst[2] has even gone so far as to construct an orthogonal basis, though any reasonable combinations should help significantly. Turek[16] has independently implemented these ideas and finds that his code is much improved over the original approach of Mead and Papanicolaou.[1]

III. INTERATOMIC POTENTIALS VIA THE MAXIMUM ENTROPY PRINCIPLE

The study of defects and structural energetics in metals is greatly aided by the concept of effective interatomic potentials. This field has suffered from the lack of uniform methods for obtaining such potentials. One of us has recently shown[18] that calculating interatomic potentials can be formulated as a problem of incomplete information: Given knowledge of the changes in the two point density-density correlation function in a condensed matter system, how can one best guess the associated energy changes? To answer this question, one must obtain guesses for higher order correlation functions (*i.e.*, triplet, four-body ...). Knowledge of these functions is necessary for calculating total energies since these depend on clusters of 3,4, ··· particles. Maxent can be used to estimate these functions, and thus produce a rigorous foundation for future work in the area. The method produces expressions for the effective potential as a functional of the pair density-density

correlation function. For further details of this approach, we refer the reader to Ref. 15.

ACKNOWLEDGEMENTS

We are very grateful to Professor John Skilling for his generous support in making this presentation possible. We would also like to thank Professor E. T. Jaynes for many stimulating discussions. One of us (DAD) would like to thank Professor Jaynes for giving the most lucid and thought-provoking lectures he has ever encountered, and for numerous helpful hints in the course of research. This research was supported in part by NSF Grant No. DMR–88–02160, and DOE Grant No. DE–FG02–84ER45130.

BIBLIOGRAPHY

1. Lawrence R. Mead and N. Papanicolaou, J. Math. Phys. **25**, 2404 (1984).
2. G. L. Bretthorst, Proceedings of the 7th Annual conference on Maximum Entropy (to be published).
3. D. A. Drabold (unpublished).
4. D. A. Drabold, Phys. Rev B **37**, 565 (1988).
5. See for example Albert Messiah, *Quantum Mechanics*, (Wiley, 1959).
6. A. Abragam, *Principles of Nuclear Magnetism,* (Oxford, 1985).
7. S. J. Knak Jensen and E. Kjaersgaard Hansen, Phys. Rev. B **7**, 2910 (1973).
8. P. A. Fedders and A. E. Carlsson , Phys. Rev. B **32**, 229 (1985).
9. See for example C. W. Myles and P. A. Fedders, Phys. Rev. B **9**, 4872 (1974).
10. D. A. Drabold and P. A. Fedders, Phys. Rev. B **37**, 3440 (1988).
11. D. A. Drabold and P. A. Fedders, submitted to Phys. Rev. B.
12. A. E. Carlsson and P. A. Fedders, Phys. Rev. B **34**, 3567 (1986).
13. A. E. Carlsson and P. A. Fedders (unpublished).
14. R. H. Brown and A. E. Carlsson, Phys. Rev. B **32**, 6125 (1985).
15. S. Glanville, A. T. Paxton and M. W. Finnis, J. Phys. F **18**, 693 (1988).
16. I. Turek, J. Phys. C **21**, 3251 (1988).
17. C. M. Nex, J. Phys. A **11**, 653 (1978).
18. A. E. Carlsson , Phys. Rev. Lett. **59**, 1108 (1987).