*On sabbatical leave from Argonne National Laboratory, Argonne, Ill.
[1]N. Cabrera, V. Celli, and R. Manson, Phys. Rev. Lett. 22, 346 (1969).
[2]G. Benedek, Phys. Rev. Lett. 35, 234 (1974).
[3]B. R. Williams, J. Chem. Phys. 55, 1315, 3220 (1971).
[4]B. F. Mason and B. R. Williams, J. Chem. Phys. 61, 2765 (1974).
[5]G. Benedek and G. Seriani, Jpn. J. Appl. Phys., Suppl. No. 2, Pt. 2, 545 (1974).
[6]G. Boato, P. Cantini, and L. Mattera, Jpn. J. Appl. Phys., Suppl. No. 2, Pt. 2, 553 (1974); P. Cantini, G. P. Felcher, and R. Taterek, to be published.
[7]R. Frisch and O. Stern, Z. Phys. 84, 430 (1933).
[8]B. Wood, B. F. Mason, and B. R. Williams, J. Chem. Phys. 61, 1435 (1974).
[9]H. U. Finzel, H. Frank, H. Hoinkes, M. Luschka, H. Nahr, H. Wilsch, and U. Wonka, Surf. Sci. 49, 577 (1975).
[10]J. A. Meyers and D. R. Frankl, Surf. Sci. 51, 61 (1975).
[11]G. Boato, P. Cantini, and L. Mattera, Surf. Sci. 55, 141 (1976).

New Variational Method with Applications to Disordered Systems

J. M. Luttinger*
The Rockefeller University, New York, New York 10021
(Received 29 June 1976)

A novel variational procedure for the partition function based on a generalization of the "coherent state" representation is presented. As an example of its use, I apply it to a simple model of a disordered system. It is shown to yield results which are exact in both the high and low temperature limits, the latter being particularly difficult to obtain by conventional methods.

The purpose of this brief note is to describe a novel variational method based on a simple generalization of the usual "coherent state" representation. This method seems to be well suited to those cases where there is symmetry breaking (which is hard to deal with by perturbative methods). An example—to be discussed more fully below—is the study of the density of electronic energy levels in disordered systems. Here the system is, on the average, translationally invariant, but the low-lying states are strongly localized no matter how weak the disordering potential is.

In its simplest form the method is as follows. Suppose one has a single spinless particle with a Hamiltonian $\mathcal{H}(\vec{p},\vec{r})$, where $\vec{p}$ and $\vec{r}$ are, respectively, the momentum and position operators of the particle. The partition function $Z(\beta)$ is defined by

$$Z(\beta)=\mathrm{Tr}[\exp(-\beta\mathcal{H})]. \tag{1}$$

To express this trace in a more convenient form, I introduce the set of functions $\psi(\vec{r};\vec{P},\vec{Q})$ defined by

$$\psi(\vec{r};\vec{P},\vec{Q})=U(\vec{P},\vec{Q})\chi(\vec{r}), \tag{2}$$

$$U(\vec{P},\vec{Q})\equiv\exp[+i(\vec{P}\cdot\vec{r}-\vec{Q}\cdot\vec{p})]; \tag{3}$$

$\vec{P}$ and $\vec{Q}$ are c-number vectors, and χ is the normalized function

$$(\chi,\chi)\equiv\int_{\text{all space}}[\chi(r)]^2d^3r=1. \tag{4}$$

I may also write[1] ($\hbar=1$)

$$U(\vec{P},\vec{Q})=e^{-i\vec{P}\cdot\vec{Q}/2}e^{i\vec{P}\cdot\vec{r}}e^{-i\vec{Q}\cdot\vec{p}} \tag{5}$$

and

$$\psi(\vec{r};\vec{P},\vec{Q})=e^{-i\vec{P}\cdot\vec{Q}/2}e^{i\vec{P}\cdot\vec{r}}\chi(\vec{r}-\vec{Q}), \tag{6}$$

since $\exp(-i\vec{Q}\cdot\vec{p})$ is a displacement operator.

The set of functions $\psi(\vec{r};\vec{P},\vec{Q})$ is over-complete and does not form an orthonormal set. In particular, if χ is the normalized ground-state wave function of a simple harmonic oscillator, the set $\psi(\vec{r};\vec{P},\vec{Q})$ forms the basis of the well-known "coherent state" representation widely used in quantum optics.[2] For the coherent state representation it is known[2] that the trace of any reasonable operator B is given by

$$\mathrm{Tr}(B)=\frac{1}{(2\pi)^3}\int d^3P\,d^3Q\,(\psi(\vec{P},\vec{Q}),B\psi(\vec{P},\vec{Q}))\equiv T. \tag{7}$$

I now show that (7) is still valid for ψ given by (2) with *arbitrary* normalized χ. Consider the right-

hand side of (7). By (6) I may write this as

$$T=(2\pi)^{-3}\int d^3P\,d^3Q\int d^3r\,d^3r'\,e^{-i\vec{P}\cdot\vec{r}}\chi^*(\vec{r}-\vec{Q})\langle\vec{r}|B|\vec{r}'\rangle\,e^{i\vec{P}\cdot\vec{r}'}\chi(\vec{r}'-\vec{Q}), \tag{8}$$

where $\langle\vec{r}|B|\vec{r}'\rangle$ is the matrix element of B in the coordinate representation. Carrying out first the $\vec{P}$, then the $\vec{r}'$ integration, one has

$$T=\int d^3Q\int d^3r|\chi(\vec{r}-\vec{Q})|^2\langle\vec{r}|B|\vec{r}\,\rangle. \tag{9}$$

Now, carrying out the $\vec{Q}$ integration and writing $\vec{Q}'=\vec{r}-\vec{Q}$, I find

$$T=\int d^3r\langle\vec{r}|B|\vec{r}\rangle\int d^3Q'|\chi(\vec{Q}')|^2=\int d^3r\langle\vec{r}|B|\vec{r}\rangle$$

which is just the trace in the coordinate representation.

I turn (7) with $B=\exp(-\beta\mathcal{H})$ into a variational method as follows

$$\begin{aligned}Z(\beta)=\mathrm{Tr}[\exp(-\beta\mathcal{H})]&=(2\pi)^{-3}\int d^3P\,d^3Q(\psi(\vec{P},\vec{Q}),\exp(-\beta\mathcal{H})\psi(\vec{P},\vec{Q})),\\&=(2\pi)^{-3}\int d^3P\,d^3Q(\chi U^\dagger(\vec{P},\vec{Q}),\exp(-\beta\mathcal{H})U(\vec{P},\vec{Q})\chi),\\&=(2\pi)^{-3}\int d^3P\,d^3Q(\chi,\exp[-\beta\mathcal{H}(\vec{p}+\vec{P},\vec{r}+\vec{Q})]\chi),\end{aligned} \tag{10}$$

where I have used the elementary results $U^\dagger(\vec{P},\vec{Q})\vec{p}U(\vec{P},\vec{Q})=\vec{p}+\vec{P}$ and $U^\dagger(\vec{P},\vec{Q})\vec{r}U(\vec{P},\vec{Q})=\vec{r}+\vec{Q}$.

For any reasonable operator A, I have

$$(\chi,\exp(A)\chi)\geq\exp(\chi,A\chi). \tag{11}$$

This is a minor variation on the well-known Jensen inequality for convex functions, as is seen by going into the representation φ_a which diagonalizes A,

$$(\chi,\exp(A)\chi)\equiv\textstyle\sum_a|(\chi,\varphi_a)|^2e^a\equiv\sum_a p_a e^a, \tag{12}$$

$$p_a\equiv|(\chi,\varphi_a)|^2\geq 0 \text{ and } \textstyle\sum_a p_a=\sum_a|(\chi,\varphi_a)|^2=(\chi,\chi)=1. \tag{13}$$

By Jensen's inequality, since e^x is convex, $\sum_a p_a e^a\geq\exp(\sum_a p_a a)=\exp(\chi,A\chi)$ which is just (11). Therefore[3]

$$Z(\beta)\geq(2\pi)^{-3}\int d^3P\,d^3Q\exp[-\beta(\chi,\mathcal{H}(\vec{p}+\vec{P},\vec{r}+\vec{Q})\chi)]. \tag{14}$$

I now use as the criterion for choosing the "best" χ the condition that this inequality be as strong as possible. That is, I choose χ so that the right-hand side of (14) is as large as possible. How well this works, and in what situations, can only be determined by experience.

As an example of the use of this method, consider a particle interacting with a collection of randomly distributed scattering centers

$$\mathcal{H}=\frac{\vec{p}^2}{2}+u_{\mathrm{wall}}(\vec{r})+\sum_{j=1}^{N}u(\vec{r}-\vec{R}_j)\equiv\frac{\vec{p}^2}{2}+v(\vec{r}), \tag{15}$$

where $m=1$, $u_{\mathrm{wall}}(\vec{r})$ is the potential which confines the particle to the volume V but is zero in V, $u(\vec{r})$ is a short-ranged potential without bound states, and $\vec{R}_j$ are the positions of the N scattering centers. The partition function will now depend on $\vec{R}_j$, but will fluctuate negligibly about its mean value $\langle Z\rangle$ given by

$$\langle Z\rangle=\int_V\prod_j\frac{d^3R_j}{V}Z. \tag{16}$$

Once I have $\langle Z\rangle$, I may obtain the density of states per unit volume $[g(\epsilon)]$ by inverting

$$\int_0^\infty g(\epsilon)e^{-\beta\epsilon}d\epsilon=\lim_{\substack{V\to\infty\\N\to\infty\\N/V\to\rho}}\frac{1}{V}\langle Z\rangle. \tag{17}$$

To compute the right-hand side of (14), I note first

$$\mathcal{H}(\vec{p}+\vec{P},\vec{r}+\vec{Q})=\tfrac{1}{2}(\vec{p}+\vec{P})^2+v(\vec{r}+\vec{Q}) \tag{18}$$

and

$$(\chi,\mathcal{H}(\vec{p}+\vec{P},\vec{r}+\vec{Q})\chi)=\tfrac{1}{2}(\chi,\vec{p}^2\chi)+(\chi,\vec{p}\chi)\cdot\vec{P}+\tfrac{1}{2}\vec{P}^2+(\chi,v(\vec{r}+\vec{Q})\chi). \tag{19}$$

The $\vec{P}$ integration in (14) can be done at once giving

$$Z(\beta)\geq(2\pi\beta)^{-3/2}\int d^3Q\,\exp[-\beta(\chi,v(\vec{r}+\vec{Q})\chi)]\exp\{-\tfrac{1}{2}\beta[(\chi,\vec{p}^2\chi)-(\chi,\vec{p}\chi)^2]\}. \tag{20}$$

It is easy to see that for a large system ($V\to\infty$) the effect of the wall potential is simply to limit $\vec{Q}$ to V, so (20) becomes

$$Z(\beta)\geq(2\pi\beta)^{-3/2}\int_V d^3Q\,\{\prod_{j=1}^{N}\exp[-\beta\int d^3r|\chi|^2u(\vec{r}+\vec{Q}-\vec{R}_j)]\}\exp\{-\tfrac{1}{2}\beta[(\chi,\vec{p}^2\chi)-(\chi,\vec{p}\chi)^2]\}. \tag{21}$$

Averaging over the $\vec{R}_j$ is elementary, and I find[4] (ρ is the density of scattering centers)

$$\langle Z\rangle\geq(2\pi\beta)^{-3/2}\int_V d^3Q\,\exp\{-\rho\int_{\text{all space}}d^3R_1[1-\exp(-\beta\int d^3r|\chi|^2u(\vec{r}+\vec{Q}-\vec{R}_1))]\}$$
$$\times\exp\{-\tfrac{1}{2}\beta[(\chi,\vec{p}^2\chi)-(\chi,\vec{p}\chi)^2]\}. \tag{22}$$

By writing $\vec{R}_1=\vec{R}+\vec{Q}$, I note that the integrand becomes independent of $\vec{Q}$ (this is the average translational invariance of the problem), and I finally obtain

$$\langle Z\rangle/V=(2\pi\beta)^{-3/2}\exp\Big(-\beta\{\tfrac{1}{2}[(\chi,\vec{p}^2\chi)-(\chi,\vec{p}\chi)^2]+\rho\int d^3R[1-\exp(-\beta\int|\chi|^2u(\vec{r}-\vec{R})\,d^3r)]\}\Big)$$
$$\equiv(2\pi\beta)^{-3/2}\exp(-\beta C[\chi]). \tag{23}$$

The "best" value of χ is therefore obtained by minimizing $C[\chi]$ subject to $(\chi,\chi)=1$.

To slightly simplify the discussion, assume χ to be real so that $(\chi,\vec{p}\chi)=0$ and

$$C[\chi]=(\chi,\tfrac{1}{2}\vec{p}^2\chi)+(\rho/\beta)\int d^3R\,(1-e^{-\beta\epsilon(\vec{R})}), \tag{24}$$

where

$$\epsilon(\vec{R})=\int d^3r\,\chi^2(\vec{r})u(\vec{r}-\vec{R}). \tag{25}$$

The condition that $C[\chi]$ be stationary subject to $(\chi,\chi)=1$ yields

$$\tfrac{1}{2}\vec{p}^2\chi+\varphi\chi=E\chi, \tag{26}$$

$$\varphi(\vec{r})\equiv\rho\int d^3R\,e^{-\beta\epsilon(\vec{R})}u(\vec{r}-\vec{R}). \tag{27}$$

These equations (24)–(27) are identical to those found by Friedberg and myself[4] using path integral techniques, under the assumption that β was very large and the potential u was very weak. As was shown in that paper, this result is precise enough to give the leading term of $\langle Z\rangle/V$ for large β [and therefore $g(\epsilon)$ for small ϵ] exactly. This is the well-known Lifshitz result which comes physically from localized states (electrons trapped in large regions free of scattering centers), and which is far from trivial to be obtained directly from the original Schrödinger equation. The first correction has also been calculated exactly[4] for short-ranged u, the result depending only on the scattering length a of the potential u. If I use Eqs. (24)–(27) and assume that

$$a'\equiv(2\pi)^{-1}\int u(\vec{r})\,d^3r \tag{28}$$

exists, I find that the correction is identical in form to the exact result with a replaced by a'. Using the inequality $a\leq a'$, it is easy to see that

$$\langle Z\rangle_{\text{variational}}\leq\langle Z\rangle$$

as is required.

There is more. For small β, I may expand the exponent in (24) and obtain

$$C=(\chi,\tfrac{1}{2}\vec{p}^2\chi)+2\pi\rho a'. \tag{29}$$

Clearly $(\chi,\frac{1}{2}\vec{p}^2\chi)$ may be made as small as I please by taking χ very slowly varying, and therefore

$$C_{\text{minimum}}=2\pi\rho a'. \tag{30}$$

This gives from (23)

$$\langle Z\rangle_{\text{variational}}=Z_0e^{-\beta(2\pi\rho a')}, \tag{31}$$

where Z_0 is the value of Z in the absence of scattering centers. But the right-hand side of (31) is exactly what I would obtain if all the unperturbed levels were just shifted by $2\pi\rho a'=\rho\int d^3r\,u(\vec{r})$, which is the result of first-order perturbation theory. Small β corresponds to high temperatures (and therefore large particle energy) where perturbation theory is valid, so this variational method gives the first two terms of $Z(\beta)$ for small β correctly. It is therefore accurate at both large and small β and I might expect it to give a reasonable approximation over the entire range of β.[5]

This work was carried out during my tenure as a Guggenheim Fellow at the Rockefeller University. I would like to thank the Guggenheim Foun-

dation for its generous support and my colleagues at Rockefeller University for their warm hospitality.

*Permanent address: Department of Physics, Columbia University, New York, N. Y. 10027. This research was supported in part by the National Science Foundation.

[1]I have used the identity $\exp(A_1+A_2)=\exp[\frac{1}{2}(A_2,A_1)]\times\exp(A_1)\exp(A_2)$, valid if (A_2,A_1) is a c number.

[2]For a review of the properties of these functions see R. J. Glauber, *Fundamental Problems in Statistical Mechanics II*, edited by E. G. D. Cohen (North-Holland, Amsterdam, 1968), p. 140 ff. See also J. R. Klauder and E. C. G. Sudarshan, *Fundamentals of Quantum Optics* (Benjamin, New York, 1968), p. 105 ff.

[3]A bound of this type for magnetic systems has been given previously: E. Lieb, Commun. Math. Phys. 31, 327 (1973). The functions used were the magnetic analogs of the coherent states, there being no variable element analogous to χ. Therefore Lieb has a lower bound and not a variational method. Lieb also has a related upper bound to $Z(\beta)$, which one can also turn into a variational method. In the examples I have looked at, this proved rather weak, and I shall not discuss it here.

[4]The procedure is identical to that used by R. Friedberg and J. M. Luttinger, Phys. Rev. B 12, 4460 (1975), for an analogous average.

[5]A numerical investigation of (24)–(27) for all β is underway at present in collaboration with Mr. C. Y. Yu.

Spin Waves for Single Electrons in Paramagnetic Metals

A. Janossy* and P. Monod
Laboratoire de Physiques des Solides,† Université Paris-Sud, 91405-Orsay, France
(Received 9 February 1976)

A new type of wave propagation in pure metals in a magnetic field is reported. This mode is analogous to cyclotron waves but is carried by the spin of conduction electrons. Oscillations in the microwave transmission spectra of copper and gold foils with ferromagnetic layers deposited onto them are attributed to these waves.

Several types of cyclotron waves[1] are known in pure metals placed in a magnetic field. These are due to the phase coherence of the orbital motion of conduction electrons on particular parts of the Fermi surface. In this Letter we describe waves analogous to cyclotron waves which give rise to Gantmakher-Kaner oscillations[2] but where the phase coherence is maintained by the Larmor precession of the spins of electrons traveling along the magnetic field. We call this type of propagation *Larmor waves*.

We will first briefly explain how Larmor waves can appear in the absence of Fermi-liquid effects[3] following the work of Dyson[4] but in the long-mean-free-path limit where the diffusion model breaks down. We then give an account of their observation in copper and gold.

We consider the transmission of an electromagnetic wave due to the phase coherence of conduction-electron spins through a metal slab in a static magnetic field that is tilted away from the surface normal by an angle α. We assume a pure metal at low temperatures where the mean free path is of the order of the sample thickness l and the skin depth δ is much smaller. Since electrons cross the skin region of the excited surface in a time much shorter than the Larmor period $2\pi/\omega_L$ they develop a small transverse magnetization perpendicular to both the static and the instantaneous direction of the rf field of frequency $\omega\neq\omega_L$. The polarization of a group of electrons with constant velocity component v_H along the static magnetic field propagates into the metal, precessing with the Larmor frequency. Thus a circularly polarized exciting field $h_0e^{-i\omega t}$ gives rise to a Larmor wave of wave vector $(\omega-\omega_L)/v_H$. The transverse magnetization at a distance z from the exciting surface is oscillating as a function of the static magnetic field[5]:

$$M(z)=Bv_H^{-1}\exp\left|\frac{i(\omega-\omega_L)-\tau^{-1}}{v_H\cos\alpha}z-i\omega t\right|, \qquad (1)$$

where B depends on the details of the excitation and τ is the momentum relaxation rate. The transmitted field—the observed quantity—is proportional to the magnetization $M(l)$. In general the distribution in v_H has to be taken into account. Larmor waves are nonresonant modes; they are sensitive to collisions and to destructive interference between electrons following different paths as well. We can show that the resulting Larmor wavelength corresponds to electrons hav-