Thermodynamics of magnetic response

1996 J. Phys. A: Math. Gen. 29 1411

(http://iopscience.iop.org/0305-4470/29/7/013)

J. Phys. A: Math. Gen. **29** (1996) 1411–1419. Printed in the UK

Thermodynamics of magnetic response

A M Stewart†

Department of Applied Mathematics, Research School of Physical Sciences and Engineering, The Australian National University, Canberra ACT 0200, Australia

Received 19 May 1995

Abstract. The thermodynamic behaviour of a magnetic system coupled to external fields is studied in a situation in which the paramagnetic and diamagnetic terms are treated on an equal footing. It is shown by explicit calculation that the linear response functions satisfy certain inequalities and in particular that the paramagnetic and diamagnetic response functions satisfy *separate* inequalities. The magnetic moment is found to be gauge invariant only if both the paramagnetic and diamagnetic terms are included.

1. Introduction

If the Hamiltonian of a system in thermodynamic equilibrium contains terms involving externally applied fields, for example electric or magnetic fields, the properties of the system will change when these fields are varied. Thermodynamic quantities like the free energy will change as will statistical averages of quantum mechanical operators such as those for the magnetic and electric moments. Changes of the statistical averages that are proportional to the fields are described by susceptibilities; the changes of the thermodynamic properties are quadratic in the fields. Since the discovery of equation (1) [1], which enables the derivatives of statistical averages of systems in thermodynamic equilibrium to be expressed exactly, the description of systems in which the coupling to the external fields in the Hamiltonian is linear, such as spin paramagnetism [2], has become well established [3, 4].

However, within the context of linear response theory, less attention appears to have been paid to situations in which the coupling is nonlinear, such as with magnetic systems in which the magnetic field couples to the orbital moment so that the Hamiltonian contains terms involving $(\mathfrak{p} - e\boldsymbol{A})^2$, where $\mathfrak{p}$ is the canonical momentum of a particle and $\boldsymbol{A}$ is the electromagnetic vector potential. The terms linear in $\boldsymbol{A}$ in the expansion of the square give rise to orbital paramagnetism and those quadratic in $\boldsymbol{A}$ to diamagnetism. In many treatments it is assumed that the diamagnetic terms are negligible in comparison with the paramagnetic [4], which often occurs in practice. However, in cases when it does not it is instructive to construct the formalism for the situation in which they are both taken into account. This is done in this paper and it is found that the magnetic moment is gauge invariant only if both the paramagnetic and diamagnetic parts are included and that the paramagnetic and diamagnetic response functions satisfy separate inequalities required by thermodynamics. Some applications of the inequalities are discussed.

† E-mail: andrew.stewart@anu.edu.au

0305-4470/96/071411+09$19.50

2. Derivatives of the free energy

We start with the mathematical identity [1]

$$\frac{\partial}{\partial\lambda}\mathrm{e}^{-\beta\mathcal{H}} = -\mathrm{e}^{-\beta\mathcal{H}}\int_0^\beta \mathrm{e}^{x\mathcal{H}}\frac{\partial\mathcal{H}}{\partial\lambda}\mathrm{e}^{-x\mathcal{H}}\,\mathrm{d}x \tag{1}$$

where $\mathcal{H}(\lambda, \mu)$ is an operator that is a function of the parameters λ and μ, which are c numbers (i.e. not operators themselves), but not a function of β. $\mathcal{H}$ does not, in general, commute with its derivatives with respect to these parameters; if it does the right-hand side of equation (1) is simply $-\beta\exp(-\beta\mathcal{H})\partial\mathcal{H}/\partial\lambda$. The validity of equation (1) is verified by noting that if the sides of it are denoted as $Q(\beta)$ then they both satisfy the same equation $\partial Q/\partial\beta = -\partial/\partial\lambda\{\mathcal{H}\exp(-\beta\mathcal{H})\}$ with $Q(0) = 0$. Applications of this identity in quantum and statistical mechanics have been discussed by Wilcox [3].

We take the trace of (1). The operators on the right-hand side cycle under the trace to eliminate the exponents containing x and the integral is evaluated trivially to give $\partial Z/\partial\lambda = -\beta Z\langle\partial\mathcal{H}/\partial\lambda\rangle_T$, where the partition function is $Z = \mathrm{Tr}\{\exp(-\beta\mathcal{H})\}$, and the symbol $\langle\mathcal{O}\rangle_T$ for any operator $\mathcal{O}$ indicates the ensemble average $\mathrm{Tr}\{\exp(-\beta\mathcal{H})\mathcal{O}\}/Z$ with $\beta = 1/kT$ where T is the temperature. The Helmholtz free energy is given by $F = -kT\log Z$ and so

$$\frac{\partial F}{\partial\lambda} = \left\langle\frac{\partial\mathcal{H}}{\partial\lambda}\right\rangle_T. \tag{2}$$

The next step is to differentiate equation (2) with respect to another parameter μ. This second derivative has three terms. The first comes from the derivative of the operator itself and is simply $\langle\partial^2\mathcal{H}/\partial\mu\partial\lambda\rangle_T$. The second comes from the derivative of Z in the denominator of the ensemble average and is $-\partial Z/\partial\mu Z^{-1}\langle\partial\mathcal{H}/\partial\lambda\rangle_T$. The third term comes from differentiating the $\exp(-\beta\mathcal{H})$ in the ensemble average with respect to μ using equation (1); this gives

$$-\int_0^\beta \mathrm{d}y\left\langle\mathrm{e}^{y\mathcal{H}}\frac{\partial\mathcal{H}}{\partial\mu}\mathrm{e}^{-y\mathcal{H}}\frac{\partial\mathcal{H}}{\partial\lambda}\right\rangle_T.$$

Collecting the three terms together we obtain the second derivative of the free energy:

$$\frac{\partial^2 F}{\partial\mu\partial\lambda} = \left\langle\frac{\partial^2\mathcal{H}}{\partial\mu\partial\lambda}\right\rangle_T - \int_0^\beta \mathrm{d}y\left\langle\mathrm{e}^{y\mathcal{H}}\delta\frac{\partial\mathcal{H}}{\partial\mu}\mathrm{e}^{-y\mathcal{H}}\delta\frac{\partial\mathcal{H}}{\partial\lambda}\right\rangle_T \tag{3}$$

where $\delta\mathcal{O} = \mathcal{O} - \langle\mathcal{O}\rangle_T$, the fluctuation of the operator from its ensemble average value. For magnetostatic systems the first term will be shown to be associated with the diamagnetic and the second term with the paramagnetic susceptibility.

3. Algebra

It is useful to examine the algebraic properties of a quantity related to the second term of equation (3):

$$\chi'_{\mu\lambda} = \frac{1}{Z}\int_0^\beta \mathrm{d}y\ \mathrm{Tr}\,\mathrm{e}^{-\beta\mathcal{H}}\mathrm{e}^{y\mathcal{H}}\mathcal{O}_\mu\mathrm{e}^{-y\mathcal{H}}\mathcal{O}_\lambda \tag{4}$$

where the operators $\mathcal{O}$ are Hermitian. The operators may be cycled under the trace to give $\mathcal{O}_\mu$ at the end. If the variable of integration is changed to $x = \beta - y$ the integral becomes the same as that in equation (4) but with λ and μ interchanged. Hence χ' and the second term of equation (3) are invariant under the interchange of λ and μ. The complex conjugate of $\chi'_{\mu\lambda}$ is obtained by reversing the order of the operators under the trace and taking their adjoints. Since they are all Hermitian it follows that $\chi'^{*}_{\mu\lambda} = \chi'_{\lambda\mu} = \chi'_{\mu\lambda}$ and therefore that χ' is real.

If either of the two operators $\mathcal{O}$ commutes with $\mathcal{H}$ then the exponentials cancel and the integration may be done trivially to give $\chi'_{\mu\lambda} = \langle \mathcal{O}_\mu \mathcal{O}_\lambda \rangle_T / kT$, which is greater than or equal to zero if $\mu = \lambda$. The same expression also holds if the temperature is much greater than the upper eigenvalue of the Hamiltonian (if it is bounded) by taking only the first term in the expansion of the exponential, but the structures of the correlation functions in the numerators are different in the two cases. In the first there are only diagonal elements in the correlation function as the eigenstates of one of the operators are the same as those of the Hamiltonian; for the high temperature case there are also off-diagonal elements in the correlation function.

Next we express the trace in equation (4) as a sum over the eigenstates $|a\rangle$ with energies E_a of the operator $\mathcal{H}$ and insert a sum $\sum_b |b\rangle\langle b|$ over the complete set of these states in between the operators $\mathcal{O}_\lambda$ and $\mathcal{O}_\mu$. The integrand is $\exp\{y(E_a - E_b)\}$ which may be integrated to give

$$\chi'_{\mu\lambda} = \frac{1}{Z} \sum_{a,b} \frac{\mathrm{e}^{-\beta E_a} - \mathrm{e}^{-\beta E_b}}{E_b - E_a} \langle a|\mathcal{O}_\mu|b\rangle\langle b|\mathcal{O}_\lambda|a\rangle \tag{5}$$

where $Z = \sum_a \exp\{-\beta E_a\}$. The term containing the energies is always greater than zero because if $E_b > E_a$ the first exponential is greater than the second and *vice versa*. If, in addition, $\mu = \lambda$ then the product of the matrix elements is a perfect square and it follows that $\chi'_{\lambda\lambda} \geqslant 0$. It is shown in appendix A, by purely algebraic methods, that a further inequality is valid, namely $\chi'_{\lambda\lambda}\chi'_{\mu\mu} \geqslant \chi'^2_{\mu\lambda}$.

When the states a and b are degenerate the denominator of equation (5) diverges. However, this difficulty may be circumvented by taking the limit $(\mathrm{e}^{-\beta E_a} - \mathrm{e}^{-\beta E_b})/(E_b - E_a) \to \beta \mathrm{e}^{-\beta E_a}$ as $E_a \to E_b$. Equation (5) may then be arranged in a more familiar form because when $E_a = E_b$ the sums over the two exponents are the same with a and b interchanged (apart from interchanging μ and λ) and (5) becomes

$$\chi'_{\mu\lambda} = \frac{1}{Z} \sum_a \mathrm{e}^{-\beta E_a} \sum_b \{\delta_{E_b,E_a}/kT + 2(1 - \delta_{E_b,E_a})/(E_b - E_a)\} \,\mathrm{Re}\{\langle a|\mathcal{O}_\mu|b\rangle\langle b|\mathcal{O}_\lambda|a\rangle\}. \tag{6}$$

The purpose of the Kronecker delta is to indicate that only terms with $E_b = E_a$ are to be included in the first (Curie) term of equation (6) and only terms with $E_b \neq E_a$ in the second (the Van Vleck [5] term). In other words the factor $(1 - \delta_{E_b,E_a})/(E_b - E_a)$ is *defined* to be zero when $E_b = E_a$.

At zero temperature only the (non-degenerate) ground state $|0\rangle$ is thermally populated and so

$$\chi'_{\mu\lambda} = 2 \sum_{b\neq 0} \frac{\mathrm{Re}\{\langle 0|\mathcal{O}_\mu|b\rangle\langle b|\mathcal{O}_\lambda|0\rangle\}}{(E_b - E_0)}. \tag{7}$$

4. The second-order free energy is negative

Let the Hamiltonian operator of a system be $\mathcal{H} = \mathcal{H}_0 + \lambda\mathcal{V}$, where the eigenfunctions $|a\rangle$ and energies E_a of the *total* Hamiltonian $\mathcal{H}$ are known and $\mathcal{V}$ is an operator with strength λ that represents the interaction of the system with external fields. In cases where there is more than one field the number λ is taken to multiply all of them. The derivative $\partial\mathcal{H}/\partial\lambda = \mathcal{V}$. From equation (2) $\partial F/\partial\lambda = \langle\mathcal{V}\rangle_T$ where the ensemble averages are to be performed with the full Hamiltonian. The second derivative $\partial^2 F/\partial\lambda^2$ is given by the second term only of equation (3) because $\partial^2\mathcal{H}/\partial\lambda^2 = 0$. By expanding the free energy in a Taylor series $F(\lambda+\delta\lambda) = F(\lambda) + \delta\lambda\partial F/\partial\lambda + (\partial^2 F/\partial\lambda^2)\delta\lambda^2/2 + \cdots$ we get

$$F(\lambda+\delta\lambda) = F(\lambda) + \langle\mathcal{V}\rangle_T\delta\lambda - \frac{\delta\lambda^2}{2}\int_0^\beta \mathrm{d}y\, \langle \mathrm{e}^{y\mathcal{H}}\delta\mathcal{V}\mathrm{e}^{-y\mathcal{H}}\delta\mathcal{V}\rangle_T + \cdots \tag{8a}$$

which may be expressed alternatively as

$$F(\lambda+\delta\lambda) = F(\lambda) + \langle\mathcal{V}\rangle_T\delta\lambda - \frac{\delta\lambda^2}{2Z}\sum_{a,b}\frac{\mathrm{e}^{-\beta E_a} - \mathrm{e}^{-\beta E_b}}{E_b - E_a}|\langle b|\delta\mathcal{V}|a\rangle|^2 + \cdots. \tag{8b}$$

Although the change in the free energy that is first order in $\mathcal{V}$ may be of either sign, the second-order term is always negative, a result obtained by Peierls [6]. The result does not depend upon λ being a small quantity.

5. Derivatives of the Hamiltonian

Consider a system of particles of charge e and mass m, which are assumed for simplicity to be the same, situated in a uniform externally applied magnetic field $\boldsymbol{B}(\mathfrak{r})$ which arises from the vector potential $\boldsymbol{A}(\mathfrak{r}, R) = \boldsymbol{B}\times(\mathfrak{r}-\boldsymbol{R})/2$, where $\boldsymbol{R}$ is the origin of the vector potential. A change of $\boldsymbol{R}$ corresponds to making a gauge transformation [7]. The Hamiltonian operator for one particle is

$$\mathcal{H}^1 = (\mathfrak{p} - e\boldsymbol{A})^2/2m - \frac{e}{2m}\boldsymbol{B}\cdot 2\boldsymbol{s} - \sum_\sigma R_\sigma\mathfrak{o}_\sigma + e\phi(\mathfrak{r}) \tag{9}$$

where $\mathfrak{p} = m\mathfrak{v} + e\boldsymbol{A}$ is the canonical momentum, $\mathfrak{v}$ is the velocity vector operator, $\mathfrak{r}$ the position operator and $\boldsymbol{s}$ is the spin. The R_σ are fields that couple to general operators $\mathfrak{o}_\sigma$ of the system, for example for an electric field $\boldsymbol{E}$ this term would be $-\boldsymbol{E}\cdot\mathfrak{p}$ where $\mathfrak{p} = e\mathfrak{r}$, and ϕ is a one particle potential. The total Hamiltonian consists of the sum of the single-particle Hamiltonians plus the interaction between the particles $\mathcal{H}_{int}$ which does not depend explicitly on the applied fields. The expansion of the orbital term gives two extra terms, $e^2\boldsymbol{A}^2/2m$ and, from the result that div $\boldsymbol{A} = \boldsymbol{0}$, $-\boldsymbol{A}\cdot\mathfrak{p}e/m = -\boldsymbol{B}\cdot\mathfrak{l}e/2m$, where the canonical orbital angular momentum $\mathfrak{l} = (\mathfrak{r}-\boldsymbol{R})\times\mathfrak{p}$. The derivatives of the Hamiltonian with respect to the fields R_s is $\partial\mathcal{H}/\partial R_\sigma = -\mathcal{O}_\sigma$, where $\mathcal{O}_\sigma = \sum_i \mathfrak{o}^i_\sigma$, the sum of the single-particle operators, for example the gradient of $\mathcal{H}$ with respect to the electric field would be $\nabla_{\boldsymbol{E}}\mathcal{H} = -\sum_i \mathfrak{p}_i$, where $\nabla_{\boldsymbol{E}} = \hat{x}\partial/\partial E_x + \hat{y}\partial/\partial E_y + \hat{z}\partial/\partial E_z$. The gradient with respect to the vector magnetic field $\boldsymbol{B}$ is $-\nabla_{\boldsymbol{B}}\mathcal{H} = \mathcal{M} = \sum_i \mathfrak{m}^i$, where the single-particle operator $\mathfrak{m}$ is

$$\mathfrak{m} = \frac{e}{2m}(\mathfrak{l} + 2\mathfrak{s}) - \frac{e^2}{4m}(\mathfrak{r}-\boldsymbol{R})\times(\boldsymbol{B}\times(\mathfrak{r}-\boldsymbol{R})). \tag{10}$$

This expression may be verified by using the relation $\mathcal{H}^1(\boldsymbol{A}+\delta\boldsymbol{A}) - \mathcal{H}^1(\boldsymbol{A}) = -e\delta\boldsymbol{A}\cdot(\mathfrak{v} - \delta\boldsymbol{A}e/2m)$, the order of the operators being important and the chain rule for differentiation.

Equation (10) may be expressed in the form $\mathfrak{m} = \{(\mathfrak{r}-\boldsymbol{R})\times m\mathfrak{v} + 2\mathfrak{s}\}e/2m$. But $(\mathfrak{r}-\boldsymbol{R})\times\mathfrak{v}e/2$ is simply the classical operator for the magnetic moment due to orbital motion about point $\boldsymbol{R}$ ($\mathfrak{s}e/m$ is that for the spin), and consequently we identify $\mathcal{M}$ as the operator for total magnetic moment [5]. A significant feature of equation (10) is that its expectation value is gauge invariant. The operator $(\mathfrak{p} - e\boldsymbol{A})$ is well known to be gauge invariant [7]; the expectation value of the $\boldsymbol{R}$ dependent term $-\boldsymbol{R}\times\langle v\rangle e/2$ is zero because $\langle v\rangle$ is zero, a non-zero transport current being incompatible, by Maxwell's equations, with a uniform magnetic field. Accordingly the expectation value of $\mathfrak{m}$ is independent of $\boldsymbol{R}$, a change of which amounts to a change of gauge. Commonly, only the first term of equation (10) is used as the magnetic moment operator, but we see that it is necessary to use both terms to maintain gauge invariance. We note also that it is *not* possible to express the magnetic part of the Hamiltonian as $-\mathcal{M}\cdot\boldsymbol{B}$ because then the diamagnetic term would not be given correctly.

The only second derivative of the Hamiltonian that is non-zero is that with respect to the magnetic field. By using Cartesian components of the vector $\mathfrak{r}' = \mathfrak{r} - \boldsymbol{R}$ and the identity $\mathfrak{r}' \times (\boldsymbol{B} \times \mathfrak{r}') = \boldsymbol{B}(\mathfrak{r}'^2) - \mathfrak{r}'(\mathfrak{r}' \cdot \boldsymbol{B})$ we write the second derivative in matrix form:

$$\frac{-\partial^2\mathcal{H}}{\partial B_\mu \partial B_\lambda} = \frac{\partial \mathcal{M}_\lambda}{\partial B_\mu} = -\frac{e^2}{4m}\sum_i \begin{bmatrix} y'^2 + z'^2 & -x'y' & -x'z' \\ -x'y' & z'^2 + x'^2 & -y'z' \\ -x'z' & -y'z' & x'^2 + y'^2 \end{bmatrix} \tag{11}$$

where λ and μ range through x, y, z starting at the top left corner of the matrix and the sum is over the i particles. All higher derivatives are zero.

It is necessary to say why it is possible to ignore the effect of gauge transformations, specifically the dependence on the origin of the vector potential, in conventional treatments of magnetism, i.e. those that consider only the first term of equation (10). It has been shown elsewhere [8] that if the origin of vector potential is shifted by a distance $\boldsymbol{R}$ then the paramagnetic and diamagnetic moments, those coming from the first and second terms of equation (10), respectively, have a contribution quadratic in $\boldsymbol{R}$ added and subtracted from them. The sum, of course, remains constant. It follows that if the total magnetic moment on one atomic site is calculated correctly by taking the origin of the vector potential to be at its nucleus then it will have the same magnetic moment whatever its position in the lattice with respect to the origin of $\boldsymbol{R}$.

6. Expansion of the free energy

As we have obtained explicit forms for the derivatives of the Hamiltonian we may expand the free energy as a Taylor series in all the applied fields R_σ as:

$$F(R_\sigma + \delta R_\sigma, B + \delta B, \text{etc}) = F(R_\sigma, B, \text{etc}) + \delta F_1 + \delta F_2 + \cdots$$

where $\delta F_1 = \sum_\lambda \delta\lambda(\partial F/\partial\lambda)$ and $\delta F_2 = \sum_{\mu\lambda} \delta\mu\delta\lambda(\partial^2 F/\partial\mu\partial\lambda)/2$, where μ and λ represent the various fields. First we select from δF_1 all those terms that are linear in the fields. These are

$$-\sum_\sigma \delta R_\sigma \langle \mathcal{O}_\sigma \rangle_T - \delta\boldsymbol{B}\cdot\langle \mathcal{L} + 2\mathcal{S}\rangle_T e/2m.$$

There is left over a term

$$+\delta\boldsymbol{B}\cdot\left\langle \sum_i \mathfrak{r}'_i \times (\boldsymbol{B} \times \mathfrak{r}'_i)\right\rangle_T e^2/4m$$

that is bilinear in $\boldsymbol{B}$ which arises from the second term of equation (10). By explicitly calculating the triple cross product in the ensemble average it is found that this may be written as

$$\left\langle\left(\sum_i \mathfrak{r}'_i \times \{\boldsymbol{B} \times \mathfrak{r}'_i\}\right)_\lambda\right\rangle_T e^2/4m = \sum_\mu \chi^{\mathrm{D}}_{\lambda\mu} B_\mu$$

where the matrix of the coefficients is

$$\chi^{\mathrm{D}}_{\mu\lambda} = -\frac{e^2}{4m}\left\langle \sum_i \begin{bmatrix} y'^2 + z'^2 & -x'y' & -x'z' \\ -x'y' & z'^2 + x'^2 & -y'z' \\ -x'z' & -y'z' & x'^2 + y'^2 \end{bmatrix}\right\rangle_T \tag{12}$$

and so the left over term is $\sum_{\mu\lambda} \chi^{\mathrm{D}}_{\mu\lambda} B_\mu \delta B_\lambda$.

The terms in δF_2 that are bilinear in the fields may be written as

$$\delta F_2 = -\tfrac{1}{2} \sum_{\mu,\lambda} \{\chi_{\mu\lambda} \delta R_\mu \delta R_\lambda + \chi^{\mathrm{D}}_{\mu\lambda} \delta B_\mu \delta B_\lambda\}. \qquad (13)$$

The first term in this expression comes from the second term of equation (3) where

$$\chi_{\mu\lambda} = \int_0^\beta \mathrm{d}y \, \langle \mathrm{e}^{y\mathcal{H}} \delta \mathcal{O}_\mu \mathrm{e}^{-y\mathcal{H}} \delta \mathcal{O}_\lambda \rangle_T. \qquad (14)$$

R_λ stands for B_λ or a general field, and $\mathcal{O}_\lambda = \partial\mathcal{H}/\partial R_\lambda$ is the operator that corresponds to it. $\mathcal{O}_\lambda$ may be $\mathcal{O}_\sigma$ or $(\mathcal{L} + 2\mathcal{S})$; the diamagnetic second term in equation (10) will give rise to terms that are of higher order in the fields and so do not appear here. The second term in equation (13) comes from the first term of equation (10). Together with the diamagnetic term that comes from δF_1 it is found that the diamagnetic contribution to the free energy has the simple form $-\sum_{\mu\lambda} \chi^{\mathrm{D}}_{\mu\lambda} B_\mu B_\lambda / 2$ (see appendix B). Accordingly, to bilinear order in the fields, the change of free energy is given by

$$F(R_\mu + \delta R_\mu, R_\lambda + \delta R_\lambda) - F(R_\mu, R_\lambda) = -\sum_\lambda \delta R_\lambda \langle \mathcal{O}_\lambda \rangle_T - \sum_{\mu\lambda} \chi_{\mu\lambda} \delta R_\mu \delta R_\lambda / 2 + \delta F^{\mathrm{D}} \qquad (15)$$

where the change of free energy due to diamagnetism δF^{D} is obtained from the expression $F^{\mathrm{D}} = -\sum_{\mu\lambda} \chi^{\mathrm{D}}_{\mu\lambda} B_\mu B_\lambda / 2$. In the case where the $\langle \mathcal{O}_\lambda \rangle_T$ are linearly proportional to the fields the free energy is given in appendix B as simply $F = -\sum_{\mu\lambda} \chi_{\mu\lambda} R_\mu R_\lambda / 2 - \sum_{\mu\lambda} \chi^{\mathrm{D}}_{\mu\lambda} B_\mu B_\lambda / 2$, where the R_μ include all the fields including the magnetic fields B_μ but the second term includes only the magnetic fields. With only magnetic fields present the change in free energy will be proportional to $(\chi_{\mu\lambda} + \chi^{\mathrm{D}}_{\mu\lambda})$, where the first term is the paramagnetic term, and will, therefore, as shown in the next section, be independent of gauge.

7. Susceptibilities

The change of the expectation value of an operator $\mathcal{O}_\lambda$ in response to the change of a parameter μ of the Hamiltonian is described by the susceptibility $\chi_{\mu\lambda} = \partial/\partial\mu \langle \mathcal{O}_\lambda \rangle_T$. The differentiation may be performed in the same way as in obtaining equation (3) to give

$$\chi_{\mu\lambda} = \int_0^\beta \mathrm{d}y \left\langle \mathrm{e}^{y\mathcal{H}} \delta \left(-\frac{\partial\mathcal{H}}{\partial\mu} \right) \mathrm{e}^{-y\mathcal{H}} \delta \mathcal{O}_\lambda \right\rangle_T. \qquad (16)$$

Since $-\partial\mathcal{H}/\partial\mu$ is given by the corresponding operator $\mathcal{O}_\mu$ it is seen that the coefficients $\chi_{\mu\lambda}$ in the expansion for the free energies in equation (13) are identical to the susceptibilities $\partial/\partial R_\mu \langle \mathcal{O}_\lambda \rangle_T$. As these susceptibilities have the algebraic form shown in equation (4) they have the properties derived previously of being real, symmetric in μ and λ, of having the diagonal elements greater than zero and of obeying the inequality $\chi_{\lambda\lambda}\chi_{\mu\mu} \geqslant \chi^2_{\mu\lambda}$.

The expectation value of the magnetic moment operator of equation (10) may be differentiated with respect to B_μ to give

$$\frac{\partial \langle \mathcal{M}_\lambda \rangle_T}{\partial B_\mu} = \frac{e}{2m} \frac{\partial \langle \mathcal{L}_\lambda + 2\mathcal{S}_\lambda \rangle_T}{\partial B_\mu} - \frac{e^2}{4m} \frac{\partial}{\partial B_\mu} \left\langle \sum_i (\mathfrak{r}'_i \times \{\boldsymbol{B} \times \mathfrak{r}'_i\})_\lambda \right\rangle_T \qquad (17)$$

or $\partial \langle \mathcal{M}_\lambda \rangle_T / \partial B_\mu = \chi_{\mu\lambda} + \chi^{\mathrm{D}}_{\mu\lambda}$ the sum of the paramagnetic and diamagnetic susceptibilities respectively. Since the magnetic moment of equation (10) is gauge invariant, the total susceptibility is gauge invariant too; the present proof of this result is more direct than those given previously [5, 7] as it does not depend on any particular wavefunction basis set or perturbation scheme. A cross term such as the magnetoelectric susceptibility is given by replacing ∂B_μ by ∂E_μ, it also has a paramagnetic and diamagnetic part, but only the paramagnetic part is involved in the inequalities of appendix A.

8. Inequalities

The diamagnetic contribution to the free energy $F^{\mathrm{D}} = -\sum_{\mu\lambda} \chi^{\mathrm{D}}_{\mu\lambda} B_\mu B_\lambda/2$ of equation (15) is a *positive* definite bilinear form because it arises from the term $e^2\boldsymbol{A}^2/2m$ in the Hamiltonian which is always positive as $\boldsymbol{A}$ is real. Consider the contribution to this free energy $-\chi^{\mathrm{D}}_{\mu\mu} B^2_\mu/2$ that arises from the presence of only one field B_μ. If this is to be positive it is required that the diagonal elements of the diamagnetic susceptibility must be negative; equation (12) shows that this is the case. Consider next the contribution to the diamagnetic free energy that arises from two fields B_μ and B_λ which is $-(\chi^{\mathrm{D}}_{\mu\mu} B^2_\mu + \chi^{\mathrm{D}}_{\lambda\lambda} B^2_\lambda + 2\chi^{\mathrm{D}}_{\mu\lambda} B_\mu B_\lambda)/2$. The two cross coefficients are equal because they come from the second partial derivatives of the Hamiltonian. This contribution to the free energy may be written in the algebraically identical form

$$\frac{1}{2}\left\{\left(\sqrt{-\chi^{\mathrm{D}}_{\mu\mu}}\,B_\mu - \frac{\chi^{\mathrm{D}}_{\mu\lambda}}{\sqrt{-\chi^{\mathrm{D}}_{\mu\mu}}}\,B_\lambda\right)^2 - B^2_\lambda\left(\chi^{\mathrm{D}}_{\lambda\lambda} - \chi^{D2}_{\mu\lambda}/\chi^{\mathrm{D}}_{\mu\mu}\right)\right\}.$$

For this to be positive definite it is necessary, noting that the diagonal elements are negative, for $\chi^{\mathrm{D}}_{\lambda\lambda}\chi^{\mathrm{D}}_{\mu\mu} \geqslant \chi^{D2}_{\mu\lambda}$. We demonstrate that the elements of the tensor have this property. Consider, for example, $\chi^{\mathrm{D}}_{xx}\chi^{\mathrm{D}}_{yy} - \chi^{D2}_{xy}$. From equation (12) this is proportional to $(\langle z^2\rangle_T\langle r^2\rangle_T + \langle x^2\rangle_T\langle y^2\rangle_T - \langle xy\rangle^2_T)$. But $\langle x^2\rangle\langle y^2\rangle \geqslant \langle xy\rangle^2$ by the Schwarz–Cauchy inequality [9] so therefore the inequality involving the two susceptibilities is satisfied.

Next we consider the change in free energy that is due to those bilinear terms that are not associated with diamagnetism. From equations (8) the second-order term in the free energy is negative, therefore the term $\sum_{\mu\lambda} \chi_{\mu\lambda}\delta R_\mu \delta R_\lambda$ in equation (15) must be a *positive* definite bilinear form. Accordingly, by the arguments used above, thermodynamics requires that $\chi_{\lambda\lambda} \geqslant 0$ and $\chi_{\lambda\lambda}\chi_{\mu\mu} \geqslant \chi^2_{\mu\lambda}$. However, this has already been shown to be true for functions having the form of equation (16) by the algebraic arguments in appendix A so it is verified that the requirements of thermodynamics are satisfied. We see that the paramagnetic and diamagnetic susceptibilities satisfy separate inequalities.

9. Applications of the inequalities

The inequality $\chi_{\lambda\lambda}\chi_{\mu\mu} \geqslant \chi^2_{\mu\lambda}$ may be applied to the tensor properties of crystals. When the paraelectric tensor, for example, is referred to the principal axes the off-diagonal components vanish [10, 11], but when it is not referred to the principal axes the inequality gives restrictions upon the measured components that must be satisfied. The inequality gives an upper limit upon the magnetoelectric susceptibility in terms of the paraelectric and paramagnetic susceptibilities. Brown *et al* [12] found that the magnetoelectric susceptibility of known magnetoelectric substances is only 1% of the maximum permitted value. On the other hand for a piezoelectric material, such as Rochelle salt, the ratio of the piezoelectric coupling term to the square root of the product of the diagonal responses is 70% and for the magnetostrictive material Permindur it is 24% [13], so an order of magnitude improvement in the cross response coefficients of such materials may not be expected.

Another application is the checking of microscopic theories that describe the physical properties of materials. One such example is a mean field theory that describes the magnetic susceptibility of metallic samarium materials on the basis of the interactions between the ions that are the result of the indirect exchange interaction that is mediated by conduction electrons. We merely sketch the argument here, greater detail is given elsewhere [14, 15].

In the tripositive rare earth ion samarium the orbital L and spin S magnetic moments of the 4f shell of the ion do not have the same proportionality to each other at different temperatures because of the admixture of different multiplet levels. The mean field equations that have been proposed to describe the linear response in the paramagnetic regime are [14]

$$-\langle L+2S\rangle_T/\mu_B = A_{MM}B + A_{MS}2B_{ex} \qquad (18a)$$

$$-\langle S\rangle_T/\mu_B = A_{SM}B + A_{SS}2B_{ex} \qquad (18b)$$

where B is the applied magnetic field and the exchange field is given by

$$B_{ex} = -\lambda\langle S\rangle_T/\mu_B + B\alpha/2.$$

The A parameters are susceptibilities that describe the response of the ion to these fields; they are temperature dependent, going to zero at high temperature, and depend only on the energy levels and wavefunctions of the ion, including crystal field effects. Because A_{MM} describes the response of the ionic magnetization to the magnetic field which couples to it, it is a diagonal susceptibility and is positive, as is A_{SS} because the exchange field couples to the spin. The quantity A_{SM} is an off-diagonal susceptibility and therefore satisfies the inequality $A_{MM}A_{SS} \geqslant A_{SM}^2$. The sum of the ionic magnetization and that induced in the conduction electrons is $-\mu_B\{\langle L+2S\rangle_T + \alpha\langle S\rangle_T\} = \chi B$ which defines the total susceptibility χ (apart from the Pauli and diamagnetic terms) [14]. The quantities λ and α are molecular field parameters which in principle may have any value and μ_B is the Bohr magneton. It is straightforward to show that the total susceptibility given by this model is

$$\frac{\chi}{\mu_B^2} = A_{MM} - \frac{A_{MS}^2}{A_{SS}} + \frac{(A_{MS}+\alpha A_{SS})^2}{A_{SS}(1-2\lambda A_{SS})}. \qquad (19)$$

It is clear that this total susceptibility is always positive in the temperature region where the expression for it is applicable. The sum of the first two terms is positive as shown above, the numerator of the third is a perfect square, $A_{SS}(T)$ is always positive and so the third term is always positive if $1 \geqslant 2\lambda A_{SS}(T)$. This defines a temperature below which the susceptibility diverges and a phase transition occurs. The interacting susceptibility is always positive above this temperature whatever the values of the molecular field parameters and in this respect is valid thermodynamically.

Acknowledgments

It is a pleasure to acknowledge useful discussions with Dr K Kumar and also with Professor K Golden during the early stages of this work.

Appendix A

From equation (5) the quantity χ' is given by

$$\chi'_{rs} = \sum_{a,b} f(a,b)\langle a|\mathcal{R}|b\rangle\langle b|\mathcal{S}|a\rangle = \sum_{a,b} f(a,b)(R'_{ab} + \mathrm{i}R''_{ab})(S'_{ba} + \mathrm{i}S''_{ba})$$

where $f(a,b) = (\mathrm{e}^{-\beta E_a} - \mathrm{e}^{-\beta E_b})/(E_b - E_a)Z$ and consequently $f(a,b) = f(b,a)$, and $f(a,b) \geqslant 0$. From the Hermiticity of the operators $\mathcal{R}$ and $\mathcal{S}$ it follows that $R'_{ab} = R'_{ba}$ and $R''_{ab} = -R''_{ba}$. The imaginary part of χ' vanishes because the terms with (a,b) and (b,a) cancel, therefore

$$\chi'_{\mathrm{rs}} = \sum_{a,b} f(a,b)(R'_{ab}S'_{ab} + R''_{ab}S''_{ab}) \tag{A1}$$

whence it follows that $\chi'_{\mathrm{rs}} = \chi'_{sr}$ and $\chi'_{rr} \geqslant 0$. We next prove that $\chi'_{rr}\chi'_{ss} \geqslant \chi^2_{\mathrm{rs}}$. This is equivalent to proving that the quantity $Y \geqslant 0$ where $Y = \chi'_{rr}\chi'_{ss} - \chi'^2_{\mathrm{rs}}$ or

$$Y = \sum_{abcd} f(a,b)f(c,d)\{(R'^2_{ab} + R''^2_{ab})(S'^2_{cd} + S''^2_{cd}) - (R'_{ab}S'_{ab} + R''_{ab}S''_{ab})(R'_{cd}S'_{cd} + R''_{cd}S''_{cd})\}. \tag{A2}$$

We note that the terms that exchange a with b or c with d are equal. Instead of (A2) we take half the sum of the terms with indices in the orders $(abcd)$ and $(cdab)$ which gives

$$Y = \sum_{abcd} f(a,b)f(c,d)\{(R'^2_{ab} + R''^2_{ab})(S'^2_{cd} + S''^2_{cd}) + (S'^2_{ab} + S''^2_{ab})(R'^2_{cd} + R''^2_{cd}) \\ -2(R'_{ab}S'_{ab} + R''_{ab}S''_{ab})(R'_{cd}S'_{cd} + R''_{cd}S''_{cd})\}/2. \tag{A3}$$

It is readily verified that the terms in the curly brackets are equal to

$$(R'_{ab}S'_{cd} - R'_{cd}S'_{ab})^2 + (R''_{ab}S''_{cd} - R''_{cd}S''_{ab})^2 + (R'_{ab}S''_{cd} - R''_{cd}S'_{ab})^2 + (R''_{ab}S'_{cd} - R'_{cd}S''_{ab})^2.$$

Since this is the sum of perfect squares it follows that $Y \geqslant 0$ and the inequality is proved.

Appendix B

If a quantity $F(B_\mu, B_\lambda)$ has the purely bilinear form (with no linear terms) $F = -\sum_{\mu\lambda} \chi_{\mu\lambda} B_\mu B_\lambda/2$ it follows that $\partial F/\partial\lambda = -\sum_\mu \chi_{\mu\lambda} B_\mu$ and $\partial^2 F/\partial\mu\partial\lambda = -\chi_{\mu\lambda}$. Conversely if the relationship between the first and second derivatives of a function is of this nature it follows that the function may be expressed in a purely bilinear form to this order.

References

[1] Kumar K 1965 *J. Math. Phys.* **6** 1928
[2] Kubo R 1959 *Lectures in Theoretical Physics* ed W E Brittin and L G Dunham (London: Interscience)
[3] Wilcox R M 1967 *J. Math. Phys.* **8** 962
[4] Grandy W T 1987 *Foundations of Statistical Mechanics* (Dordrecht: Reidel)
[5] Van Vleck J H 1932 *The Theory of Electric and Magnetic Susceptibilities* (London: Oxford University Press)
[6] Peierls R E 1933 *Z. Phys.* **80** 763
[7] Griffith J S 1961 *Theory of Transition Metal Ions* (Cambridge: Cambridge University Press)
[8] Stewart A M 1996 *Austral. J. Phys.* to be published
[9] Hardy G H, Littlewood J E and Polya G 1952 *Inequalities* 2nd edn (Cambridge: Cambridge University Press)
[10] Nye J F 1957 *Physics Properties of Crystals* (Oxford: Oxford University Press)
[11] Stewart A M 1994 *Austral. J. Phys.* **47** 129
[12] Brown W F, Hornreich R M and Shtrikman S 1968 *Phys. Rev.* **168** 574
[13] Mason W P 1972 Properties of transducer materials *American Institute of Physics Handbook* 3rd edn (New York: McGraw-Hill) pp 3–118
[14] Stewart A M 1993 *Phys. Rev.* B **47** 11 242
[15] Stewart A M 1993 *Rare Earths '93 Processing and Utilisation Workshop Papers* Australian Mineral Foundation, Adelaide 129. ISBN 0 908039 56 5