

A New Hierarchy of Upper and Lower Bounds on Expectation Values

U. Brandt and J. Stolze

Institut für Physik der Universität, Dortmund, Federal Republic of Germany

Received April 8, 1981

Upper and lower bounds are constructed for expectation values of functions of a real random variable with derivatives up to order $N+1$ which are alternately negative and positive over the whole range of interest. The bounds are given by quadrature formulas with weights and abscissas determined by the first $N+1$ moments of the underlying probability distribution. Application to a simple disordered phonon system yields sharp bounds on the specific heat.

I. Introduction

In view of the subtleties and uncertainties of perturbation expansions of several kinds, inequalities have proven to be a valuable tool for reliable calculations in statistical mechanics, as documented e.g. in the review of Girardeau and Mazo [1]. With some notable exceptions, however, (e.g. GKS-type bounds on correlation functions of ferromagnetic Ising spin systems [2] or Fisher's bounds on derivatives [3]) most of the statistical-mechanical inequalities derived up to now have provided bounds to partition functions or free energies but not to directly accessible physical quantities. In this contribution we wish to derive a hierarchy of (upper and lower) bounds to expectation values of a certain class of functions. These inequalities may be viewed as generalizations of Jensen's famous inequality which gives a lower bound to the expectation value of a function, provided the latter is convex and the zeroth and first moments of the underlying probability distribution are known. We generalize Jensen's inequality to the case where all moments up to a certain order are known and all derivatives up to a certain order of the function under consideration have alternately positive and negative signs over the whole interval of interest. Our bounds are given in terms of quadrature formulas the abscissas and weights of which are determined by a system of nonlinear equations involving only the given moments. The existence of bounds of this type is shown

in Sect. 2 by applying a modified Lagrange parameter method for solving constrained extremal problems. We then show how to find the best possible bounds of this type (provided they exist). In Sect. 3 we apply our method to the calculation of bounds on the specific heat of a simple (one-dim.) disordered system of masses and springs. By choosing a special complete system of states for the calculation of the traces involved, we are able to give analytically the moments of the "distribution of the squared frequencies" for each state. We then apply the formalism derived formerly to each state separately and finally perform the sum over all states. The bounds on the specific heat obtained by this method prove to be quite sharp, the lower bounds going to zero linearly for low temperatures and the upper bounds tending to a finite but small limiting value as $T \rightarrow 0$. We conclude by discussing the applicability of our method to other problems of interest.

Two appendices serve to illustrate the essential features of the above-mentioned modified Lagrange parameter method and to determine the properties of a certain function occurring in the calculation of the specific heat of oscillator systems.

II. The Hierarchy of Bounds

In this section we derive upper and lower bounds on expectation values of certain functions of a real ran-

dom variable x . These bounds may be viewed as generalizations of the well-known Jensen inequality

$$\langle f(x) \rangle \geq f(\langle x \rangle) \quad (1)$$

for convex functions f . In (1) the zeroth and first moments of the probability density governing the distribution of x are used to bound the expectation value of $f(x)$; we wish to point out how the knowledge of some further moments may be exploited to yield a series of increasingly sharp upper and lower bounds on $\langle f(x) \rangle$ if $f(x)$ and its derivatives $f^{(n)}(x)$ fulfill certain conditions.

Let us consider a function $f(x)$ which is defined on an interval $I := [a, b]$ and has the property

$$(-1)^n f^{(n)}(x) \geq 0 \quad n=0, 1, \dots, N+1 \quad (2)$$

for all $x \in I$. We are interested in bounds on the expectation value of $f(x)$

$$\langle f(x) \rangle := \int_a^b dx \rho(x) f(x) \quad (3)$$

with respect to a probability density $\rho(x)$, given the moments

$$m_v := \langle x^v \rangle \quad v=0, 1, \dots, N. \quad (4)$$

Surely we have

$$\sup_{w(x)} \int_a^b dx w(x) f(x) \geq \langle f(x) \rangle \geq \inf_{w(x)} \int_a^b dx w(x) f(x) \quad (5)$$

where the functions $w(x)$ are assumed to fulfill the conditions

$$\int_a^b dx w(x) x^v = m_v \quad v=0, 1, \dots, N \quad (6)$$

and

$$w(x) \geq 0 \quad \text{for } x \in I. \quad (7)$$

The calculation of the sup and inf in (5) is the solution of a constrained extremal problem. The constraints (6) may be accounted for by Lagrange parameters as usual; for the treatment of inequality constraints such as (7) there exists a modified Lagrange parameter method which has been applied to physical problems earlier [4, 5].* Instead of searching for extremal values of the functional $\int_a^b dx w(x) f(x)$ subject to the constraints (6) and (7)

we have to search for extremal values of the auxiliary functional

$$\begin{aligned} \Phi[w(x)] := & \int_a^b dx w(x) f(x) \\ & + \sum_{v=0}^N \alpha_v \left(\int_a^b dx w(x) x^v - m_v \right) + \int_a^b dx w(x) \lambda(x) \end{aligned}$$

with Lagrange parameters α_v and a "Lagrange function" $\lambda(x)$. The function $\lambda(x)$ is non-negative with $\lambda(x)=0$ for all x where $w(x)>0$ in the "sup" case. In the "inf" case $\lambda(x)$ is non-positive in an analogous way. Taking the functional derivative of Φ with respect to $w(x)$ we obtain the extremal condition

$$f(x) + \sum_{v=0}^N \alpha_v x^v + \lambda(x) = 0.$$

For x corresponding to non-vanishing $w(x)$ we have

$$F(x) := f(x) + \sum_{v=0}^N \alpha_v x^v = 0. \quad (8)$$

If we assume temporarily that in addition to (2) we also have $f^{(N+1)}(x) \neq 0$ we see that the solutions of (8) are isolated points* and thus the extremal probability densities are combinations of δ -functions located at these points. If we call the abscissas of the δ -functions x_i and their corresponding weights w_i we have

$$w(x) = \sum_i w_i \delta(x - x_i) \quad (9)$$

and thus

$$m_v = \sum_i w_i x_i^v \quad v=0, 1, \dots, N. \quad (10)$$

Note that (sets of) abscissas and weights may be determined from the non-linear equations (10) alone without any recourse to properties of the function f .

Let us now ask for the number of abscissas given by (8). For definiteness we consider the sup case. From the properties of λ we know that $F(x) \leq 0$ in the interval I . Thus the zeros of $F(x)$ in the interior I_0 of I are necessarily of multiplicity two (or four, etc.). Let us assume that $F(x)$ has K zeros in I_0 and S zeros on the boundary ∂I . Then by applying Rolle's theorem the first derivative $F'(x)$ has at least $2K+S-1$ zeros in I_0 , namely the K double zeros of $F(x)$ and at least $K+S-1$ ones between the zeros of $F(x)$. Applying Rolle's theorem repeatedly, we see that $F^{(N)}(x)$ has at least $2K+S-N$ zeros in I_0 . On the other hand the equation

* For the sake of completeness we have included an appendix illustrating the modified Lagrange method on an intuitive level

* This is of course true only for *one-dimensional* random variables

$$F^{(N)}(x) = f^{(N)}(x) + N! \alpha_N = 0$$

has at most one solution in I_0 and we obtain the following restriction on the number of abscissas

$$2K + S \leq N + 1. \tag{11}$$

We now abandon the additional condition $f^{(N+1)}(x) \neq 0$ and assume that we have found solutions to the equations

$$m_v = \sum_{i=1}^{K+S} x_i^v w_i \quad v=0, 1, \dots, N \tag{12}$$

subject to the condition (11). If necessary, we add additional abscissas with weight zero in the interior I_0 of the interval such that

$$2K + S = N + 1.$$

This is of course possible only if

$$S = \begin{cases} 1 & \text{for } N \text{ even} \\ 0 \text{ or } 2 & \text{for } N \text{ odd.} \end{cases}$$

By construction, these solutions fulfill

$$\sup_{w(x)} \int_a^b dx w(x) f(x) \geq \sum_{i=1}^{K+S} f(x_i) w_i \geq \inf_{w(x)} \int_a^b dx w(x) f(x). \tag{13}$$

We shall now demonstrate that among the solutions described above we can find ones in which one of the inequality signs in (13) is replaced by an equality sign. To this end we construct an interpolating polynomial $P(x)$ of degree N with the properties

$$P(x_i) = f(x_i) \quad \text{for } x_i \in I, \\ P'(x_i) = f'(x_i) \quad \text{for } x_i \in I_0.$$

As the ‘‘quadrature formula’’ involving the abscissas x_i and weights w_i is exact for polynomials up to order N , we have

$$\Delta := \int_a^b dx w(x) (f(x) - P(x)) = \int_a^b dx w(x) f(x) - \sum_{i=1}^{K+S} P(x_i) w_i \\ = \int_a^b dx w(x) f(x) - \sum_{i=1}^{K+S} f(x_i) w_i$$

for every $w(x)$ with the given set of moments. Consider now the function

$$\varphi(t) := (f(x) - P(x)) \omega(t) - (f(t) - P(t)) \omega(x)$$

with the $(N + 1)^{\text{th}}$ order polynomial $\omega(x)$ defined by

$$\omega(x) := \prod_{x_i \in I_0} (x - x_i)^2 \prod_{x_j \in \partial I} (x - x_j). \tag{14}$$

The function $\varphi(t)$ has zeros at $t = x_i$ for all $x_i \in I$ and for $t = x$, furthermore $\varphi'(t)$ has zeros for $t = x_i \in I_0$. By applying Rolle’s theorem we conclude that $\varphi^{(N+1)}(t)$ has at least one zero, say at $t = \xi$ which leads to

$$0 = \varphi^{(N+1)}(\xi) = (f(x) - P(x)) (N + 1)! - f^{(N+1)}(\xi) \omega(x)$$

or equivalently

$$f(x) - P(x) = f^{(N+1)}(\xi) \frac{\omega(x)}{(N + 1)!}.$$

Now we have to distinguish between special cases. Let us assume that N is even so that we obtain one abscissa in ∂I , say at $x = b$. In this case $f^{(N+1)}(x)$ and $\omega(x)$ are non-positive in I by (2) and (14), respectively, and thus $\Delta \geq 0$ or equivalently

$$\sum_{i=1}^{K+S} f(x_i) w_i \leq \int_a^b dx w(x) f(x) \tag{15}$$

for every admissible $w(x)$, so that (15) remains true if we replace the r.h.s. by its infimum; but then we obtain the converse of the rightmost part of (13). Thus we have shown that the special solution discussed above yields the infimum in (13). The other special cases are treated similarly and we obtain the results summarized in the following table

	N even	N odd
inf	$S = 1$ abscissa at $x = b$	$S = 0$
sup	$S = 1$ abscissa at $x = a$	$S = 2$

To summarize, we have developed bounds on $\langle f(x) \rangle$ in terms of quadrature formulas the weights and abscissas of which are expressible in terms of the given moments. For $\rho(x) \equiv 1$, i.e. for $m_v = (b^v - a^v)/(v + 1)$ these quadrature formulas reduce to the well-known Gauss, Radau, and Lobatto formulas for $S = 0, 1, 2$, respectively [6]. The existence of connections between quadrature formulas and the ‘‘reduced moment problem’’ is of course well-known to mathematicians, cf. e.g. [7]. The possibility of constructing a sequence of upper and lower bounds in the special case $f(x) = (x + a)^{-1}$ (x and a positive) was already observed earlier [8] but only the first approximation was used. The solution of the system of nonlinear equations (12) may become rather cum-

bersome if N is large. The direct solution may however be circumvented by the use of Padé approximant techniques which allow the abscissas and weights to be expressed by the poles and residues of Padé approximants of a suitably constructed "moment-generating polynomial" [9].

For the case $a=0$, $b=2$ and $m_0=1$ (normalized probability distribution) we give explicit formulas for the weights and abscissas in terms of moments for the first orders:

Sup:

$$1. \text{ Order: } x_1=0, x_2=2, w_2=\frac{m_1}{2}, w_1=1-w_2$$

$$2. \text{ Order: } x_1=0, x_2=\frac{m_2}{m_1}, w_2=\frac{m_1^2}{m_2}, w_1=1-w_2$$

$$3. \text{ Order: } x_1=0, x_2=\frac{2m_2-m_3}{2m_1-m_2}, x_3=2$$

$$w_2=(2m_1-m_2)/(2x_2-x_2^2)$$

$$w_3=(m_1x_2-m_2)/(2x_2-4)$$

$$w_1=1-w_3-w_2$$

$$4. \text{ Order: } x_1=0, x_{2,3}=-y_1 \pm \sqrt{y_1^2-y_2}$$

$$\text{with } y_1=(m_1m_4-m_2m_3)/2(m_2^2-m_1m_3)$$

$$y_2=(m_3^2-m_2m_4)/(m_2^2-m_1m_3)$$

$$w_2=(m_2-x_3m_1)/x_2(x_2-x_3)$$

$$w_3=(m_2-x_2m_1)/x_3(x_3-x_2)$$

$$w_1=1-w_2-w_3$$

Inf.:

$$1. \text{ Order: } x_1=m_1, w_1=1$$

$$2. \text{ Order: } x_1=(2m_1-m_2)/(2-m_1), x_2=2,$$

$$w_1=(2-m_1)/(2-x_1), w_2=1-w_1$$

$$3. \text{ Order: } x_{1,2}=-y_1 \pm \sqrt{y_1^2-y_2}$$

$$\text{with } y_1=(m_3-m_1m_2)/2(m_1^2-m_2)$$

$$y_2=(m_2^2-m_1m_3)/(m_1^2-m_2)$$

$$w_1=\frac{m_1-x_2}{x_1-x_2}, w_2=1-w_1.$$

III. Application to a Disordered Phonon System

Now we use our theory to obtain rigorous upper and lower bounds for the specific heat of a disordered elastic system. The specific heat (per site) is given by

$$c=N^{-1} \sum_{\nu} f(\beta^2 \lambda_{\nu}/4) \quad (16)$$

where $f(x)=x/\sinh^2 \sqrt{x}$, N the number of atoms, λ_{ν} the eigenvalues of the dynamical matrix H , i.e. the frequencies ω_{ν} of the eigenmodes are given by $\lambda_{\nu}^{1/2}$. The summation is performed over all eigenmodes. It is shown in the Appendix A that the derivatives of $f(x)$ for $x>0$ are alternately either positive or negative, but they do not change sign as x is varied.

Expression (16) can be written as

$$c=N^{-1} \int d\lambda f(\beta^2 \lambda/4) \text{Tr } \delta(\lambda-H). \quad (17)$$

Now one may choose $N^{-1} \text{Tr } \delta(\lambda-H)=\rho(\lambda)$ for the weight function and determine its moments, but the resulting bounds would be very inaccurate*. Instead we choose a suitable unity-resolving set of normalized states $\{|q\rangle\}$ [10] (i.e. $\sum_q a_q |q\rangle \langle q|=1$ where a_q are non-negative weight factors) and obtain

$$c=N^{-1} \sum_q a_q \int d\lambda f(\beta^2 \lambda/4) \langle q | \delta(\lambda-H) | q \rangle. \quad (18)$$

Now we choose $\langle q | \delta(\lambda-H) | q \rangle$ for our weight function $\rho_q(\lambda)$ and apply our procedure; the last q -summation has to be performed directly. If we were able to choose $|q\rangle$ as the eigenstates of H our result would be exact, but for a disordered system this seems to be impossible. It should be mentioned that the configuration average can be performed on the weight function $\rho_q(\lambda)$, but in the thermodynamic limit the averaging procedure is (in most cases) unnecessary.

For the demonstration of the applicability of our theory we use a one-dimensional model. We emphasize that our theory is not at all restricted to one-dimensional situations, but the calculation of higher moments becomes more tedious in higher dimensions. Our model consists of a chain of atoms of mass m with harmonic nearest-neighbour interaction with force constants K or K' . These force constants are distributed statistically and independently at the different sites (force constant disorder - FCD), thus the dynamical matrix is of the form

$$H_{ii}=(K_{i,i+1}+K_{i,i-1})/m$$

and

$$H_{i,i+1}=H_{i+1,i}=-K_{i,i+1}/m$$

where $K_{i,i+1}$ ($=K_{i+1,i}$) (the force constant of the harmonic interaction of the i -th and $(i+1)$ -th atoms) is equal to K or K' (with the probability p and $1-p$, respectively). We regard the limit $K' \rightarrow \infty$; in this limit the one-dimensional model becomes equivalent to a mass-disordered (MD) system, because a whole cluster of atoms bound together by the strong forces acts as a single, but heavier particle.

* See 'Note Added in Proof'

In the FCD model we can restrict the trace in (17) to the subspace R of those states where the displacements of all atoms within one cluster are the same, and we can replace H by PHP , where P is the projection operator on this subspace. The spectrum of PHP is bounded from below by zero and from above by $4K/m$ for any given configuration. We can number the clusters (including "one-clusters"), starting at some arbitrary point, and $v(i)$ shall denote the number of the cluster containing the site i ; the length of this cluster is denoted by B_v , and S is the total number of clusters. Obviously we have $\langle S \rangle / N = p$, the concentration of "weak" forces, where $\langle \rangle$ denotes the configuration average. We define a set of states $|q\rangle$ in R by

$$\langle i|q\rangle = S^{-1/2} \exp(iq \sum_{l<i} e_{l,l+1}) \quad (19)$$

where

$$e_{l,l+1} = \begin{cases} 1 & \text{if } K_{l,l+1} = K \\ 0 & \text{if } K_{l,l+1} = \infty \end{cases}$$

and

$$q = 2\pi v S^{-1}; \quad -S/2 < v \leq S/2.$$

In the MD picture $|q\rangle$ would simply be a plane wave. Further we define a set $\{|\bar{q}\rangle\}$ in R by

$$\langle i|\bar{q}\rangle = \langle i|q\rangle B_{v(i)}^{-1}. \quad (20)$$

Now we have $\langle \bar{q}|q\rangle = 1$ and $\sum_q |q\rangle \langle \bar{q}| = P$, the projection operator on our subspace R . Further we have

$$PHP|q\rangle = PH|q\rangle = \varepsilon_q |\bar{q}\rangle$$

with

$$\varepsilon_q = \frac{2K}{m} (1 - \cos q).$$

Now we obtain from (18)

$$c = N^{-1} \sum_q \int d\lambda f(\beta^2 \lambda/4) w_q(\lambda) \quad (21)$$

with

$$w_q(\lambda) = \langle \bar{q} | \delta(\lambda - PHP) | q \rangle.$$

Obviously $w_q(\lambda)$ is normalized and non-zero only for $0 \leq \lambda \leq 4K/m$.

Further we see

$$\begin{aligned} \lambda w_q(\lambda) &= \langle \bar{q} | \delta(\lambda - PHP) PHP | q \rangle \\ &= \varepsilon_q \langle \bar{q} | \delta(\lambda - PHP) | \bar{q} \rangle \geq 0 \end{aligned}$$

thus $w_q(\lambda)$ is also non-negative.

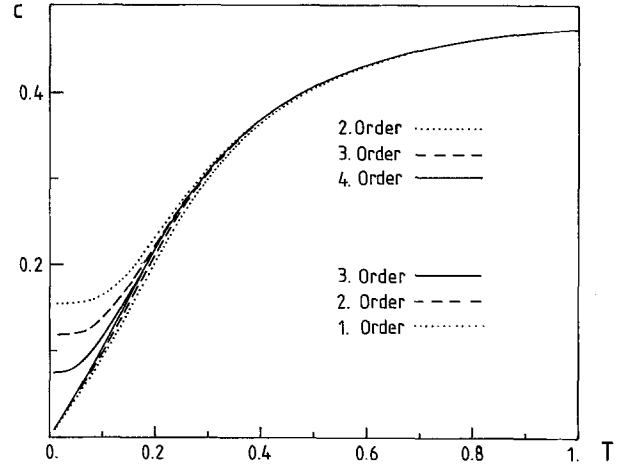


Fig. 1. Upper and lower bounds for the specific heat are plotted as functions of temperature for concentration $p=1/2$. Energy unit is $(2K/m)^{1/2}$

The first moments of w_q (averaged over all configurations) can be calculated easily:

$$m_0 = 1, \quad m_1 = \varepsilon_q g_1$$

$$m_2 = \varepsilon_q \cdot 2K m^{-1} (g_2 - g_1^2 \cos q)$$

$$m_3 = \varepsilon_q \cdot (2K m^{-1})^2$$

$$\cdot (g_3 + g_1 g_2/2 - 2 \cos q g_1 g_2 + \cos 2q g_1^3/2)$$

$$m_4 = \varepsilon_q \cdot (2K m^{-1})^3 \{g_4 + g_3 g_1 + g_2^2/2$$

$$- \cos q (2g_3 g_1 + 5g_2^2/4 + g_1^2 g_2/2)$$

$$+ \cos 2q 3g_1^2 g_2/2 - \cos 3q g_1^4/4\} \text{ etc.}$$

where $g_v = g_v(p)$ is given by the $(-v)$ th moment of the mass distribution in the MD model (in units of the elementary mass m), i.e.

$$g_v(p) = \sum_{l=1}^{\infty} \frac{(1-p)^l}{l^v} \frac{p}{1-p}.$$

We have calculated the upper bounds up to fourth order and the lower bounds up to third order for $p=1/2$. The abscissas and weights can be determined easily by hand up to this order, only the last q -integration requires some numerical effort.

In Fig. 1 we give a plot of the bounds on the specific heat (per site) as a function of temperature (we use $(2K/m)^{1/2}$ as energy unit). The upper bounds will always tend to a finite value for $T \rightarrow 0$ because the weight for the abscissa zero (the lower boundary in the λ -integral of (21)) will not vanish for all q . All lower bounds tend to zero linearly with temperature. This is true also for the asymptotic behaviour of the exact result, as one can see by a continuum limit argument for long-wavelength excitations, where the

macroscopic compressibility and density enter into the wave equation.

The slopes $\partial c/\partial T$ at $T=0$ can be calculated easily for all our lower bounds as well as for the exact solution; the results for $\frac{\partial c}{\partial T} \left(\frac{2K}{m}\right)^{1/2}$ are (numerical values are for $p=1/2$)

first order

$$p \cdot \sqrt{2/g_1} \cdot \pi/3 = 0.889$$

second order

$$p \cdot 2^{1/2} \cdot (g_1 + (g_1^2 - g_2)/2)^{-1/2} \cdot \pi/3 = 0.924$$

third order

$$p \cdot 2^{1/2} \cdot \{g_1 - (g_2 - g_1^2)/(g_3 - 3g_1 g_2/2 + g_1^3/2)\}^{-1/2} \cdot \pi/3 = 0.966$$

exact

$$(2p)^{1/2} \cdot \pi/3 = 1.047.$$

It is valuable to notice that the inclusion of the finite upper integration limit in the λ -integral of (21) (i.e. the step from an even to a successive odd order for the upper bound and from an odd to the successive even order for the lower bound) yields a considerable improvement even at low temperatures, where the contribution of this abscissa itself is exponentially small.

IV. Concluding Remarks

We started in this contribution with the discussion of a hierarchy of upper and lower bounds for the expectation value of a function having some convexity properties in connection with the incomplete moment problem. It turned out that the well known "Jensen's inequality" is just the first (and most ineffective) lower bound in the hierarchy.

We used our theory to calculate rigorous upper and lower bounds for the specific heat of a disordered harmonic system (disordered phonons). Although we chose a very simple set of states to calculate the traces (in the MD-picture simply plane waves) the results are astonishingly accurate. Our theory could be supplemented by a variational idea (for both upper and lower bounds) by choosing a set of states containing parameters (e.g. eigenstates of a parametrized effective hamiltonian) and optimizing these parameters (as in mean-field theory).

Our theory is not only applicable to disordered systems; it may be used also for the discussion of interacting systems which can be reduced to effective one-particle "hamiltonians" as e.g. in the Feynman-

polaron problem [11]. Unfortunately our method cannot be used in most cases to calculate bounds for partition functions directly for thermodynamic systems of interacting particles. In most cases we have $m_1 \sim N$ and $m_2 - m_1^2 \sim N$, thus we cannot expect to obtain corrections beyond mean-field which survive in the thermodynamic limit.

Appendix A

This appendix provides the proof that the function

$$f(x) := \frac{x}{\sinh^2 \sqrt{x}} \quad (\text{A1})$$

enjoys the property

$$(-1)^n f^{(n)}(x) \geq 0 \quad x > 0, n \geq 0. \quad (\text{A2})$$

For $n=0$ the result is obvious, for higher n we shall express $f^{(n)}(x)$ as the n -th derivative of the Laplace transform of a positive function. To this end we observe that

$$g(x) := \frac{\coth \sqrt{x}}{\sqrt{x}} = \sum_{n=-\infty}^{\infty} \frac{1}{x + n^2 \pi^2} \\ = \int_0^{\infty} e^{-xt} \sum_{n=-\infty}^{\infty} e^{-n^2 \pi^2 t} dt = \mathcal{L}[\vartheta_3(0|i\pi t)] \quad (\text{A3})$$

where we have used the symbol \mathcal{L} for Laplace transformation and

$$\vartheta_3(z|\tau) := \sum_{n=-\infty}^{\infty} e^{i\pi n^2 \tau} \cos 2n z \quad (\text{A4})$$

is one of Jacobi's theta functions [12]. From (A3) we conclude that for $n \geq 1$

$$\frac{d^n}{dx^n} f(x) = \frac{d^n}{dx^n} (-x g(x) - 2x^2 g'(x)) \\ = \mathcal{L} \left[(-t)^n \left\{ 2t \frac{d^2}{dt^2} \vartheta_3(0|i\pi t) + 3 \frac{d}{dt} \vartheta_3(0|i\pi t) \right\} \right]. \quad (\text{A5})$$

It remains to be shown that the function

$$\varphi(t) := 2t \frac{d^2}{dt^2} \vartheta_3(0|i\pi t) + 3 \frac{d}{dt} \vartheta_3(0|i\pi t) \quad (\text{A6})$$

is positive. Using (A4) we obtain

$$\varphi(t) = 2 \sum_{n=1}^{\infty} n^2 \pi^2 e^{-n^2 \pi^2 t} (2t n^2 \pi^2 - 3) \quad (\text{A7})$$

which is obviously non-negative for $t \geq 3/2\pi^2 \simeq 0.152$.

Now we apply "Jacobi's imaginary transformation" [12]

$$\vartheta_3(z|\tau) = (-i\tau)^{-1/2} \exp(z^2/i\pi\tau) \vartheta_3(z\tau^{-1} | -\tau^{-1})$$

to (A4) in order to obtain an alternative representation for $\varphi(t)$:

$$\varphi(t) = \pi^{-1/2} t^{-7/2} \sum_{n=-\infty}^{\infty} n^2 e^{-\frac{n^2}{t}} (2n^2 - 3t)$$

which is obviously non-negative for $t \leq 2/3$, such that $\varphi(t)$ is a non-negative function. This completes our proof.

Appendix B

In this appendix we wish to illustrate the essence of the Lagrange parameter method for inequality constraints. To do this we first recall the usual method for equality constraints.

Suppose we have to maximize a function $F(x, y)$ subject to the constraint

$$h(x, y) = \eta. \quad (\text{B1})$$

To do this we maximize the auxiliary function

$$L(x, y) := F(x, y) + \lambda(h(x, y) - \eta) \quad (\text{B2})$$

with respect to x , y and λ independently.

Of course the location (x_0, y_0) of the maximum as well as the maximum value $F(x_0, y_0)$ depend on η ; in fact it is easy to show that

$$\frac{dF(x_0, y_0)}{d\eta} = -\lambda \quad (\text{B3})$$

i.e. the Lagrange parameter describes the rate of change of the maximum value with respect to a change of the prescribed value of the constraint function.

Let us now consider an inequality constraint of the form

$$h(x, y) \geq \eta \quad (\text{B4})$$

describing a region (or a union of regions) of the x - y plane. If the maximum of F is situated in the interior of this region it is an unconstrained local maximum of F and we need not pay attention to (B4); the maximum value does not change if η increases by an infinitesimal amount.

If, on the other hand, the maximum is situated on the boundary, (B4) is equivalent to an equality con-

straint which has to be incorporated by a Lagrange parameter. We note that this Lagrange parameter has to be non-negative, as the maximum value of F is forced to decrease if η increases and the region shrinks.

We have thus demonstrated that a maximum problem with an inequality constraint of the form (B4) may be treated by maximizing an auxiliary function of the form (B2) with

$$\lambda = 0 \quad \text{if } h(x, y) > \eta$$

and

$$\lambda \geq 0 \quad \text{if } h(x, y) = \eta.$$

The method illustrated above can be generalized to the case of several (or even infinitely many) constraints. For a more thorough mathematical discussion of the method the reader is referred to [13].

References

1. Girardeau, M.D., Mazo, R.M.: *Adv. Chem. Phys.* **24**, 187 (1973)
2. Griffiths, R.B.: In: *Phase Transitions and Critical Phenomena*. Domb, C., Green, M.S. (eds.), Vol. 1, p. 7. London: Academic Press 1972
3. Fisher, M.E.: *J. Chem. Phys.* **42**, 3852 (1965)
4. Einhorn, M.B., Blankenbecler, R.: *Ann. Phys.* **67**, 480 (1971)
5. Yndurain, F., Yndurain, F.J.: *J. Phys.* **C8**, 434 (1975)
6. Abramowitz, M., Stegun, I.A. (eds.): *Handbook of Mathematical Functions*. New York: Dover Publications 1970
7. Shohat, J.A., Tamarkin, J.D.: *The Problem of Moments*. American Mathematical Society, Providence, Rhode Island 1950
8. Devreese, J.T.: In: *Path Integrals and Their Applications in Quantum, Statistical and Solid State Physics*. Papadopoulos, G.J., Devreese, J.T. (eds.), p. 315. New York: Plenum 1978
9. Yndurain, F.J.: In: *Padé Approximants*. Graves-Morris, P.R. (ed.), p. 45. London: The Institute of Physics 1973
10. Leschke, H., Stolze, J., Moraweck, M.: *J. Phys.* **A13**, 1311 (1980)
11. Feynman, R.P.: *Statistical Mechanics*. Reading, Mass.: Benjamin 1972
12. Whittaker, E.T., Watson, G.N.: *A Course of Modern Analysis*. Cambridge: Cambridge University Press 1927
13. Hestenes, M.R.: *Optimization Theory - The Finite Dimensional Case*. New York: J. Wiley 1975

U. Brandt
J. Stolze
Institut für Physik
Universität Dortmund
Otto-Hahn-Strasse
Postfach 500500
D-4600 Dortmund 50
Federal Republic of Germany

Note Added in Proof

After submission of this paper we became aware of the article by J.C. Wheeler, R.G. Gordon; *J. Chem. Phys.* **51**, 5566 (1969) in which the same sequence of upper and lower bounds was used to study the specific heat and other thermal properties of an ordered harmonic face-centered-cubic lattice. These authors however used

the total density of eigenvalues of the dynamical matrix ($N^{-1} \text{Tr } \delta(H - \lambda)$) as a weight function, and thus their bounds, for a given number of moments, are not as tight as those obtainable by the method discussed here.