

## A variational method in statistical mechanics

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1972 J. Phys. A: Gen. Phys. 5 1566

(<http://iopscience.iop.org/0022-3689/5/11/004>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.72

The article was downloaded on 02/06/2010 at 04:28

Please note that [terms and conditions apply](#).

## A variational method in statistical mechanics

R PHYTHIAN

Department of Physics, University College of Swansea, Singleton Park, Swansea, UK

MS received 29 June 1972, in revised form 1 August 1972

**Abstract.** A variational method, similar to the Frenkel variational method of quantum mechanics, is derived for the time development of phase functions of a system. It is shown how this method may be used to obtain approximations for quantities of interest such as the density autocorrelation function of a fluid in equilibrium.

### 1. Introduction

Great interest has recently been shown in the problem of calculating time dependent correlation functions of quantities, such as the density, for classical fluids in thermal equilibrium. As is well known, this problem is closely related to that of the linearized relaxation of the system. A variety of different approaches appears in the literature. The generalized Langevin equation based on the projection operator methods of Zwanzig (1961) and Mori (1965), has been used by Akcasu and Duderstadt (1969) and Bixon and Zwanzig (1969), while the kinetic equation approach may be found in the papers of Ortoleva and Nelkin (1969), Lebowitz *et al* (1969) and Gross (1972).

It is the purpose of the present paper to make a preliminary exploration of an alternative method which makes use of a time dependent variational principle similar to one of quantum mechanics. So far as the author is aware, the only previous use of variational methods in this context is due to Zwanzig who has presented a method for calculating approximate eigenfunctions and eigenvalues of the Liouville operator. The connection of Zwanzig's theory with the time dependent problem is rather indirect and it would seem that a time dependent variational principle is more appropriate. Mention should also be made of the variational method formulated by Nakano (1959) for the calculation of transport coefficients in quantum statistical mechanics. Although this method can probably be reformulated for classical systems it does not appear possible to make contact in this way with the work of the authors mentioned above.

A brief outline of the problem is first given in terms of time dependent phase functions of the system, and it is shown that the differential equation describing the evolution of these phase functions can be written in variational form. This follows from a temporal step by step least squares procedure, the derivation being analogous to that which gives the Frenkel variational principle of quantum mechanics. It is also pointed out that, for the problem of calculating time dependent correlation functions, an alternative variational method may be derived. The two methods lead to the same results if suitable trial functions are used. As an example of the use of such methods, an approximation is derived by employing trial functions of single particle type, and it is shown that this leads immediately to an equation for the phase density autocorrelation function derived

previously by Akcasu and Duderstadt. A more complicated approximation, based on two particle trial functions, is also obtained which appears to be equivalent to one given by Gross.

## 2. Formulation of the problem

We consider a system consisting of a large number  $N$  of identical particles described by classical mechanics, the coordinates and momentum of the  $\alpha$ th particle being denoted by  $\mathbf{q}_\alpha, \mathbf{p}_\alpha$ . The Hamiltonian of the system is taken as

$$H = \sum_{\alpha} \frac{p_{\alpha}^2}{2m} + V(\mathbf{q}_1, \dots, \mathbf{q}_N)$$

where  $V$  is the potential energy due to interactions of the particles (not necessarily pairwise) and to steady external fields. A state of thermal equilibrium is represented by the canonical ensemble which corresponds to a phase space probability density given by  $A e^{-\beta H}$  where  $A$  is the normalization constant and  $\beta$  the inverse temperature.

A nonequilibrium state is usually represented by a time dependent probability density  $P(\mathbf{q}_1, \mathbf{p}_1; \dots, \mathbf{q}_N, \mathbf{p}_N | t)$  which will frequently be written in the abbreviated form,  $P(1 \dots N | t)$ , which satisfies

$$\frac{\partial P}{\partial t} = \mathcal{L}P$$

where  $\mathcal{L}$  is the Liouville operator

$$\mathcal{L} = -\sum_{\alpha} \left( \frac{\mathbf{p}_{\alpha}}{m} \cdot \frac{\partial}{\partial \mathbf{q}_{\alpha}} - \frac{\partial V}{\partial \mathbf{q}_{\alpha}} \cdot \frac{\partial}{\partial \mathbf{p}_{\alpha}} \right).$$

The formal solution of the equation is clearly

$$P(1 \dots N | t) = e^{t\mathcal{L}} P(1 \dots N | 0).$$

Physical quantities are represented by phase functions; typical examples are the kinetic energy  $\sum_{\alpha} p_{\alpha}^2/2m$ , the particle density  $\sum_{\alpha} \delta(\mathbf{q} - \mathbf{q}_{\alpha})$ , the momentum density  $\sum_{\alpha} \mathbf{p}_{\alpha} \delta(\mathbf{q} - \mathbf{q}_{\alpha})$ , the potential energy (for pair potentials)  $\sum_{\alpha < \beta} v(\mathbf{q}_{\alpha}, \mathbf{q}_{\beta})$ . It is seen that these are of single particle type, that is, of the form  $\sum_{\alpha} \phi(\mathbf{q}_{\alpha}, \mathbf{p}_{\alpha})$  or two particle type  $\sum_{\alpha, \beta} \phi(\mathbf{q}_{\alpha}, \mathbf{p}_{\alpha}; \mathbf{q}_{\beta}, \mathbf{p}_{\beta})$ . The numerical value at time  $t$  of a quantity represented by the phase function  $\Phi$  is then given by

$$\int d1 \dots dN \Phi(1 \dots N) P(1 \dots N | t).$$

This description resembles the Schrödinger picture of quantum theory.

An alternative description is the 'Heisenberg picture' in which the time dependence is carried by the phase functions instead of by the probability density. Suppose that the system at time 0 has definite values  $\mathbf{q}_1, \mathbf{p}_1; \dots, \mathbf{q}_N, \mathbf{p}_N$  for the coordinates of all particles. As a consequence of the equations of motion the coordinates of the particles at a later time  $t$  will be given by quantities  $\mathbf{Q}_{\alpha}(t), \mathbf{P}_{\alpha}(t)$ , ( $\alpha = 1 \dots N$ ) which are functions of  $t$  and of the initial coordinates  $\mathbf{q}_1, \mathbf{p}_1; \dots, \mathbf{q}_N, \mathbf{p}_N$ . The actual value of a quantity, represented by the phase function  $\Phi$ , at time  $t$  will then be  $\Phi(\mathbf{Q}_1(t), \mathbf{P}_1(t); \dots, \mathbf{Q}_N(t), \mathbf{P}_N(t))$ . If the  $\mathbf{Q}_{\alpha}, \mathbf{P}_{\alpha}$

are written as functions of the initial coordinates in this expression we obtain a function denoted by

$$\Phi(\mathbf{q}_1, \mathbf{p}_1; \dots; \mathbf{q}_N, \mathbf{p}_N|t) \quad \text{or} \quad \Phi(1 \dots N|t)$$

which will be referred to as a time dependent phase function. If we now imagine that the initial state of the system is not definite but is given by a probability density  $P(1 \dots N)$  it is seen that the mean value of the physical quantity at time  $t$  represented by  $\Phi$  is

$$\int d1 \dots dN \Phi(1 \dots N|t) P(1 \dots N). \quad (1)$$

Comparing this with the same quantity in the Schrödinger picture and using the anti-hermitian nature of the Liouville operator (for suitable boundary conditions) we see that

$$\Phi(1 \dots N|t) = e^{-t\mathcal{L}} \Phi(1 \dots N). \quad (2)$$

This result may also be obtained directly by evaluating the time derivatives  $\dot{\Phi}(1 \dots N|0)$ ,  $\ddot{\Phi}(1 \dots N|0)$  etc by successive application of the formula

$$\frac{\partial}{\partial t} \Phi(\mathbf{Q}_1, \mathbf{P}_1 \dots \mathbf{Q}_N, \mathbf{P}_N) = \sum_x \left( \frac{\mathbf{P}_x}{m} \cdot \frac{\partial}{\partial \mathbf{Q}_x} - \frac{\partial V}{\partial \mathbf{Q}_x} \cdot \frac{\partial}{\partial \mathbf{P}_x} \right) \Phi(\mathbf{Q}_1, \mathbf{P}_1 \dots \mathbf{Q}_N, \mathbf{P}_N).$$

From the formal expression (2) for  $\Phi(1 \dots N|t)$  it follows immediately that it satisfies the equation

$$\frac{\partial \Phi(1 \dots N|t)}{\partial t} = -\mathcal{L} \Phi(1 \dots N|t) \quad (3)$$

with the initial condition  $\Phi(1 \dots N|0) = \Phi(1 \dots N)$ . It is most important to note that  $\mathcal{L}$  is here interpreted as an operator acting on functions of the initial coordinates. This equation is thus not to be confused with the one immediately preceding it, and only applies when  $V$  does not depend explicitly on  $t$ .

This Heisenberg picture is more convenient for our purposes since it enables us to discuss correlation functions such as  $\langle \Phi(1 \dots N|t) \Psi(1 \dots N|t') \rangle$  without the necessity of introducing fictitious external potentials. (Here  $\langle \dots \rangle$  denotes a canonical ensemble average.)

### 3. Variational approach

We have seen that the problem reduces to the solution of the equation

$$\frac{\partial \Phi(1 \dots N|t)}{\partial t} = -\mathcal{L} \Phi(1 \dots N|t)$$

together with an averaging over the initial distribution. An exact solution is, of course, impossible in general and one requires a criterion for selecting approximate solutions. Such a criterion is given by the least squares method. If  $w(1 \dots N)$  denotes a suitable weighting function, then it is clear that the best approximation among functions  $\Psi(1 \dots N|t)$  of some particular form is that  $\Psi$  which minimizes

$$I = \int d1 \dots dN w(1 \dots N) \left\{ \left( \frac{\partial}{\partial t} + \mathcal{L} \right) \Psi(1 \dots N|t) \right\}^2$$

for each value of  $t$  in the interval of interest. If we apply this criterion, starting at  $t = 0$  and proceeding step by step in time, the following variational equation is obtained :

$$\int d1 \dots dN w(1 \dots N) \delta \Psi(1 \dots N|t) \left( \frac{\partial}{\partial t} + \mathcal{L} \right) \Psi(1 \dots N|t) = 0.$$

(For a more detailed discussion of this point see Corson (1951) and also, in a different context, Phythian (1972).) This equation is analogous to the Frenkel variational equation of quantum mechanics.

There remains the question of the choice of the weighting function  $w$ . It is immediately apparent that, if the trial functions  $\Psi$  are completely general, this choice is irrelevant since we can achieve  $I = 0$  for all  $w$ . For a restricted set of trial functions, which is necessary for obtaining useful approximations, it is clear that the best approximation will depend upon  $w$ . The most natural choice is to take  $w$  as the initial probability distribution, and this is why our equation has been written in a form referring to the initial coordinates. The variational equation then has the final form

$$\int d1 \dots dN P(1 \dots N) \delta \Psi(1 \dots N|t) \left( \frac{\partial}{\partial t} + \mathcal{L} \right) \Psi(1 \dots N|t) = 0. \tag{4}$$

In the present paper we shall confine our attention to a system in thermal equilibrium so that  $P$  will be given by the canonical ensemble.

For the particular problem of calculating time dependent correlation functions for a system in thermal equilibrium, an alternative variational method may be formulated. Suppose that the correlation function in question is  $\langle \Phi' \Phi(t) \rangle$ , that is  $\langle \Phi' e^{-t\mathcal{L}} \Phi \rangle$ , where  $\Phi$  and  $\Phi'$  are two phase functions. Let  $\Psi(\tau)$  and  $\Psi'(\tau)$  be two functions which satisfy the conditions  $\Psi(0) = \Phi$ ,  $\Psi'(t) = \Phi'$  but which are otherwise arbitrary. Then the quantity

$$I = \left\langle \Phi' \Psi(t) - \int_0^t d\tau \Psi'(\tau) \left( \frac{\partial}{\partial \tau} + \mathcal{L} \right) \Psi(\tau) \right\rangle \tag{5}$$

is stationary for variations of  $\Psi(\tau)$  and  $\Psi'(\tau)$  about  $\Phi(\tau)$  and  $\Phi'(\tau - t)$  respectively and the stationary value in question is the correlation function  $\langle \Phi' e^{-t\mathcal{L}} \Phi \rangle$ .

This variational principle is similar to one of quantum mechanics described by Demkov (1963) and the proof is analogous. It has the advantage that the correlation function is calculated as the stationary value of some quantity. On the other hand, the method is in general more complicated because of the presence of two trial functions. However, for suitable trial functions, the two variational methods give identical results. This is the case for the particular examples to be considered below as can easily be verified.

#### 4. Illustration of the variational approach

Consider a phase function of single particle form  $\sum_{\alpha} \phi(\mathbf{q}_{\alpha}, \mathbf{p}_{\alpha})$  or, more briefly  $\sum_{\alpha} \phi(\alpha)$ . It is clear that the corresponding time dependent phase function will not, in general, be of single particle form since application of the operators  $\mathcal{L}, \mathcal{L}^2 \dots$  gives functions of two, three ... particle form. Nevertheless let us take trial functions

$$\Psi = \sum_{\alpha} \psi(\mathbf{q}_{\alpha}, \mathbf{p}_{\alpha}|t)$$

where  $\psi$  is arbitrary except for the condition  $\psi(\mathbf{q}, \mathbf{p}|0) = \phi(\mathbf{q}, \mathbf{p})$ .

The variational equation (4) gives

$$\int d1 \dots dNP(1 \dots N) \sum_{\beta} \delta\psi(\beta|t) \left( \frac{\partial}{\partial t} + \mathcal{L} \right) \sum_{\alpha} \psi(\alpha|t) = 0$$

which becomes, using symmetry and the arbitrary nature of  $\delta\psi$ ,

$$\int d2 \dots dNP(1 \dots N) \left( \frac{\partial}{\partial t} + \mathcal{L} \right) \sum_{\alpha} \psi(\alpha|t) = 0.$$

Written out in full this becomes

$$\begin{aligned} \int d2 \dots dNP(1 \dots N) \left( \frac{\partial}{\partial t} - \frac{\mathbf{p}_1}{m} \cdot \frac{\partial}{\partial \mathbf{q}_1} \right) \psi(1|t) + \int d2 \dots dNP(1 \dots N) \frac{\partial V}{\partial \mathbf{q}_1} \cdot \frac{\partial}{\partial \mathbf{p}_1} \psi(1|t) \\ + (N-1) \int d2 \dots dNP(1 \dots N) \left( \frac{\partial}{\partial t} - \frac{\mathbf{p}_2}{m} \cdot \frac{\partial}{\partial \mathbf{q}_2} + \frac{\partial V}{\partial \mathbf{q}_2} \cdot \frac{\partial}{\partial \mathbf{p}_2} \right) \psi(2|t) = 0. \end{aligned}$$

Assuming suitable boundary conditions and using the form of  $P$  given by the canonical ensemble we see that the last two terms in the last integral cancel. The equation can be rewritten

$$\left( \frac{\partial}{\partial t} - \frac{\mathbf{p}_1}{m} \cdot \frac{\partial}{\partial \mathbf{q}_1} + \frac{\partial \mathcal{U}(1)}{\partial \mathbf{q}_1} \cdot \frac{\partial}{\partial \mathbf{p}_1} \right) \psi(1|t) + (N-1) \frac{\int d2 f(1, 2) \psi(2|t)}{f(1)} = 0. \quad (6)$$

Here we have introduced the equilibrium distribution functions

$$\begin{aligned} f(1) &= \int d2 \dots dNP(1 \dots N) \\ f(1, 2) &= \int d3 \dots dNP(1 \dots N) \end{aligned}$$

and the effective potential

$$\mathcal{U}(1) = -\frac{1}{\beta} [\ln f(1)]_{\mathbf{p}_1=0}.$$

It may be seen that, if a time dependence of the form  $e^{-i\lambda t}$  is assumed for  $\psi$ , then one obtains from (6) an equation previously derived by Zwanzig for the approximate eigenfunctions of the Liouville operator (Zwanzig 1966, equation (35)).

If we consider the case of a uniform liquid then it is easily verified that the quantity

$$\chi(1|t) = f(1)\psi(1|t) + (N-1) \int d2 f(1, 2)\psi(2|t) \quad (7)$$

has a spatial Fourier transform  $\tilde{\chi}$  which satisfies the equation

$$\left( \frac{\partial}{\partial t} + \frac{i\mathbf{p} \cdot \mathbf{k}}{m} \right) \tilde{\chi}(\mathbf{k}, \mathbf{p}) = \frac{\rho \tilde{g}(\mathbf{k})}{1 + \rho \tilde{g}(\mathbf{k})} \frac{i\mathbf{p} \cdot \mathbf{k}}{m} U(\mathbf{p}) \int d\mathbf{p}' \tilde{\chi}(\mathbf{k}, \mathbf{p}')$$

where  $\rho$  is the density  $N/V$ ,  $U(\mathbf{p}) = (\beta/\pi)^{3/2} \exp(-\beta p^2/2m)$  and  $\tilde{g}(\mathbf{k})$  is the Fourier transform of the pair distribution function  $g$  defined by

$$f(1, 2) = \frac{1}{V^2} U(\mathbf{p}_1) U(\mathbf{p}_2) g(\mathbf{q}_1 - \mathbf{q}_2).$$

If the phase function in question is the phase density defined by

$$D(\mathbf{q}, \mathbf{p}) = \frac{1}{N} \sum_{\alpha} \delta(\mathbf{q} - \mathbf{q}_{\alpha}) \delta(\mathbf{p} - \mathbf{p}_{\alpha})$$

then our approximation for the quantity  $\langle D(\mathbf{q}', \mathbf{p}'|t) D(\mathbf{q}, \mathbf{p}) \rangle$  is

$$\begin{aligned} & \int d\mathbf{1} \dots d\mathbf{N} P(1 \dots N) \sum_{\alpha} \psi(\alpha|t) \frac{1}{N} \sum_{\beta} \delta(\mathbf{q} - \mathbf{q}_{\beta}) \delta(\mathbf{p} - \mathbf{p}_{\beta}) \\ &= f(\mathbf{q}, \mathbf{p}) \psi(\mathbf{q}, \mathbf{p}|t) + (N-1) \int d2 f(\mathbf{q}, \mathbf{p}; 2) \psi(2|t) \end{aligned}$$

which is just the quantity  $\chi(\mathbf{q}, \mathbf{p}|t)$  introduced above. (The  $\psi$  in this case depends on  $\mathbf{q}'$ ,  $\mathbf{p}'$  as well because of the initial condition  $\psi(1|0) = \delta(\mathbf{q}_1 - \mathbf{q}') \delta(\mathbf{p}_1 - \mathbf{p}')$ , although this has not been indicated explicitly.) The variational method thus provides a simple derivation of equation (34) in the paper of Akcasu and Duderstadt referred to previously.

A better approximation is presumably given by using two particle trial functions

$$\Psi = \sum_{\alpha < \beta} \psi(\alpha, \beta|t).$$

The variational equation (4) now gives

$$\int d\mathbf{3} \dots d\mathbf{N} P(1 \dots N) \left( \frac{\partial}{\partial t} + \mathcal{L} \right) \sum_{\alpha < \beta} \psi(\alpha, \beta|t) = 0$$

which can eventually be reduced to the two equations

$$\begin{aligned} \chi(1, 2|t) &= f(1, 2) \psi(1, 2|t) + (N-2) \int d3 f(1, 2, 3) (\psi(1, 3|t) + \psi(2, 3|t)) \\ &+ \frac{1}{2} (N-2)(N-3) \int d3 d4 f(1, 2, 3, 4) \psi(3, 4|t) \end{aligned}$$

and

$$\begin{aligned} \frac{\partial \chi(1, 2|t)}{\partial t} &= f(1, 2) \left( \frac{\mathbf{p}_1}{m} \cdot \frac{\partial}{\partial \mathbf{q}_1} + \frac{\mathbf{p}_2}{m} \cdot \frac{\partial}{\partial \mathbf{q}_2} \right) \psi(1, 2|t) + \frac{1}{\beta} \left( \frac{\partial f(1, 2)}{\partial \mathbf{q}_1} \cdot \frac{\partial}{\partial \mathbf{p}_1} + \frac{\partial f(1, 2)}{\partial \mathbf{q}_2} \cdot \frac{\partial}{\partial \mathbf{p}_2} \right) \psi(1, 2|t) \\ &+ \frac{N-2}{\beta} \int d3 \left( \frac{\partial f(1, 2, 3)}{\partial \mathbf{q}_1} \cdot \frac{\partial \psi(1, 3|t)}{\partial \mathbf{p}_1} + \frac{\partial f(1, 2, 3)}{\partial \mathbf{q}_2} \cdot \frac{\partial \psi(2, 3|t)}{\partial \mathbf{p}_2} \right). \end{aligned}$$

This is equivalent to the equations derived by Gross by a different method. It is seen that the interparticle potentials do not appear explicitly in the equations. (Incidentally we have shown that the assumption of pair potentials is not necessary for the derivation of this result.) The approximation appears rather intractable and it is probably worthwhile examining alternative ones generated by less general two particle trial functions. This and other questions will be considered in a future paper.

### Acknowledgments

The author is grateful to a referee for drawing his attention to the work of Nakano.

**References**

- Akcasu A Z and Duderstadt J J 1969 *Phys. Rev.* **188** 479–86  
Bixon M and Zwanzig R 1969 *Phys. Rev.* **187** 267–72  
Corson E M 1951 *Perturbation Methods in the Quantum Mechanics of n-Electron Systems* (Glasgow: Blackie)  
Gross E P 1972 *Ann. Phys.* **69** 42–55  
Lebowitz J L, Percus J K and Sykes J 1969 *Phys. Rev.* **188** 487–504  
Mori H 1965 *Prog. theor. Phys.* **33** 423–55  
Nakano H 1959 *Prog. theor. Phys.* **22** 453–55  
Ortoleva P and Nelkin M 1969 *Phys. Rev.* **181** 429–37  
Phythian R 1972 *J. Fluid Mech.* **53** 469–80  
Zwanzig R 1961 *Lectures in Theoretical Physics* vol 3 ed W E Britten (New York: Wiley)  
—— 1966 *Phys. Rev.* **144** 170–7