PHYSICAL REVIEW VOLUME 100, NUMBER 2 OCTOBER 15, 1955

Statistical Mechanics of Rotating Buckets*

J. M. BLATT, S. T. BUTLER, AND M. R. SCHAFROTH
The F. B. S. Falkiner Nuclear Research and Adolph Basser Computing Laboratories, School of Physics,
The University of Sydney, Sydney, Australia

(Received April 4, 1955)

A system in temperature equilibrium is enclosed in a container and this container is set rotating. We show that the effective moment of inertia is the same as what one would compute in the usual way, for both classical and quantum statistical mechanics. This apparently trivial theorem has applications for the theories of superfluidity of liquid helium and of superconductivity of metals.

1. INTRODUCTION

No one doubts that a bucket of water, set rotating in such a way that thermal equilibrium is maintained at all times (e.g., no turbulent currents are set up), would show the conventional relation between angular momentum L and angular velocity ω:

$$L=I\omega, \tag{1.1}$$

where I, the moment of inertia, has the classical value

$$I=I_0=NM\langle(x^2+y^2)\rangle_{\text{Av}}. \tag{1.2}$$

Here N is the number of molecules, M the mass of each molecule, x and y are coordinates measured from the axis of rotation (the z-axis), and the braces denote an average over the volume of the bucket occupied by the fluid.

Yet no one, to our knowledge, has proved this theorem under general conditions, and it is the purpose of this paper to do so. Contrary to one's first impression, the proof is by no means trivially easy, and the theorem has important applications for the theory of superfluidity of liquid helium and the theory of superconductivity. These applications will be discussed in subsequent papers.

In Sec. 2, we consider a statistical system set in linear motion; the method of discussing the statistical mechanics of such a system is developed in some detail; the fact that the apparent mass of the system is equal to the actual mass NM is a trivial consequence of Galilean invariance, but our discussion allows us to derive a nontrivial equipartition theorem for the kinetic energy associated with the bulk motion of the fluid. In Sec. 3, we develop the statistical mechanics of a statistical assembly inside a rotating container, and we define the moment of inertia. Section 4 contains a proof that this moment of inertia is equal to (1.2) in classical statistical mechanics. Before we go on to quantum statistical mechanics, we develop, in Sec. 5, the concept of a "correlation length," which turns out to be of crucial importance for all the subsequent work. Sections 6 and 7 contain a proof of (1.2) in quantum statistical mechanics; this proof is more restricted than the classical one, and exceptional systems (in particular, the ideal Bose-Einstein gas below its condensation point) exist which violate (1.2). Some discussion and conclusions are given in Sec. 8, but the important applications of this theorem are given in subsequent papers.

2. LINEAR MOTION AND THE EQUIPARTITION THEOREM

We consider a box of volume V filled with an assembly of N identical particles, each of mass M. We intend to develop the formalism of statistical mechanics appropriate to discuss linear motion of the whole system with constant linear velocity v. The fact that the linear momentum contributed by the fluid is NMv is of course a trivial consequence of Galilean invariance; but our formal approach will enable us to deduce a nontrivial equipartition theorem from this fact.

Let us first consider the statistical mechanics of the box *at rest*, paying special attention to the method of including the effect of the walls of the box. The Hamiltonian of the system is

$$H=\sum_{i=1}^{N}\mathbf{p}_i^2/2M+V(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N). \tag{2.1}$$

The conventional procedure is the following: the potential energy V includes the effects of all forces between the particles of the fluid, but does *not* include the effects of forces between particles of the fluid and the walls of the box. These latter effects are taken into account approximately by replacing the actual, rough walls of the box by perfectly reflecting mirrors, or, in quantum-mechanical terms, by boundary conditions on the admissible wave functions. All wave functions must vanish whenever any one of the coordinates $\mathbf{r}_i$ lies on or outside the walls of the container. The free energy F of the system is then given by ($\beta=1/kT$)

$$\exp(-\beta F)=\text{Trace}[\exp(-\beta H)], \tag{2.2}$$

where the trace is to be taken over any complete set of wave functions which satisfy these particular boundary conditions. It is important to realize that the trace (2.2) is not even defined unless the boundary conditions on the wave functions are specified. Furthermore, in order to define thermodynamic quantities such as the free

---

* Also supported by the Nuclear Research Foundation within the University of Sydney.

energy which remain constant with time (in thermal equilibrium), it is necessary that the boundary conditions imposed on the wave functions be themselves time-independent.

Let us now return to the boundary conditions for the box at rest; what is the justification for replacing the actual rough walls of the box by perfectly reflecting mirrors? Consider a typical collision between a molecule of the fluid and the actual wall: the collision results in a reflection of the molecule back into the fluid, but this reflection is not specular. The transverse component of momentum is not preserved, because the wall is rough and can take up transverse momentum; similarly the normal component of the momentum is not merely changed in sign. Furthermore the wall itself is in heat motion, so that collision between a molecule and the wall will in general result in an exchange of energy as well as in an exchange of momentum. However, once thermal equilibrium has been established within the system the over-all, average effect of collisions between molecules and the wall is the same as if the wall were a perfectly reflecting mirror surface without heat motion. For any one wall collision in which a molecule loses energy, there is another wall collision in which some other molecule gains energy; for any one wall collision in which a molecule loses transverse momentum, another molecule gains transverse momentum by a wall collision, etc. Thus it is a good approximation, in thermal equilibrium, to replace the actual rough walls by perfectly reflecting mirror surfaces. On the other hand, this method is quite useless in considering the approach to thermal equilibrium after the wall of the box has been heated up slightly, to a temperature higher than the temperature of the fluid. For by replacing the wall by a perfect mirror, we have lost the very mechanism by which equilibrium is established in this case.

One may object: Why can one not dispense with all this, and simply include the coordinates and momenta of the molecules making up the wall into the Hamiltonian H? This procedure is contrary to the spirit of statistical mechanics, and is in principle impossible. The difficulty is merely removed one step, not overcome. The wall has an outside surface, and this outside surface interacts with its surroundings. If we do not wish to make approximations, the surroundings must now be included in the Hamiltonian as well, until we finally include the whole universe into the Hamiltonian. This is of course an impossible program. No matter how we do it, at a certain point we must make a cut between particles included in our system, and the rest of the world. This rest of the world *must* be treated schematically, in some approximation, not precisely by including it in the Hamiltonian. Once we realize that approximations of this kind are in principle necessary, there is no reason against making the simplest sensible approximation, namely the replacement of the walls by perfect mirrors.

Let us now consider the thermal equilibrium state of the fluid in a box moving with a constant linear velocity v in the x-direction. We again replace the walls of the box by perfectly reflecting mirrors, and write the effect of these mirror surfaces as boundary conditions on admissible wave functions. We observed before that, in order to do statistical mechanics, it is necessary that the boundary conditions imposed on the wave functions be independent of time. Since the boundary conditions refer to the walls of the container, the only way to fulfill this requirement is to use a set of coordinates in which the walls of the container correspond to time-independent values of the coordinates; i.e., in order to do statistical mechanics at all, we are *forced* to transform to a system of coordinates moving along with the container. This transformation is not one of convenience, which could be avoided, but is a necessary step in order to do statistical mechanics at all.

The transformation to the moving coordinate system is a canonical transformation; we obtain it most simply by writing down the Lagrangian corresponding to (2.1), and making the substitutions:

$$x_i' = x_i - vt, \quad y_i' = y_i, \quad z_i' = z_i. \tag{2.3}$$

Transforming back to the Hamiltonian formalism, we get the following expressions for the canonical momenta:

$$p_{xi}' = p_{xi}, \quad p_{yi}' = p_{yi}, \quad p_{zi}' = p_{zi}. \tag{2.4}$$

We note that the momentum variable p_{xi}' represents the linear momentum of particle number i with respect to the stationary coordinate system, *not* with respect to the moving coordinate system. The transformed Hamiltonian is

$$H' = H - v \sum_{i=1}^{N} p_{xi} = H - vP, \tag{2.5}$$

where P is the x-component of the total linear momentum of all particles. The Hamiltonian H' is *not* numerically equal to the original Hamiltonian H; this is a consequence of the fact that the transformation (2.3), (2.4) involves the time explicitly.

We now determine the linear momentum of the fluid in our moving box, in thermal equilibrium. This is given by

$$\bar{P}(v) = \frac{\mathrm{Trace}[P \exp(-\beta H')]}{\mathrm{Trace}[\exp(-\beta H')]} = \frac{\mathrm{Trace}\{P \exp[-\beta(H - vP)]\}}{\mathrm{Trace}\{\exp[-\beta(H - vP)]\}}. \tag{2.6}$$

The trace is to be taken over all wave functions of the desired symmetry (symmetric for Bose-Einstein statistics, antisymmetric for Fermi-Dirac statistics) which satisfy the boundary conditions at the walls of the box.

In classical statistical mechanics the trace represents an integral over the $6N$-dimensional phase space.

We evaluate (2.6) by making a second canonical transformation:

$$\begin{aligned} x_i''&=x_i', & y_i''&=y_i', & z_i''&=z_i', \\ p_{xi}''&=p_{xi}'-Mv, & p_{yi}''&=p_{yi}', & p_{zi}''&=p_{zi}'. \end{aligned} \tag{2.7}$$

The fact that it *is* a canonical transformation can be seen by writing down the commutators; the additional term Mv commutes with everything. When we substitute (2.7) into the effective Hamiltonian H', (2.5), we obtain the new Hamiltonian H'':

$$\begin{aligned} H'=H''&=\sum_{i=1}^{N}(\mathbf{p}_i'')^2/2M+V(\mathbf{r}_1'',\mathbf{r}_2'',\cdots,\mathbf{r}_N'')-\tfrac{1}{2}NM_v{}^2 \\ &=H(p_i'',q_i'')-\tfrac{1}{2}NMv^2, \end{aligned} \tag{2.8}$$

where now $H(p_i'',q_i'')$ is formally the same function of its variables as the initial Hamiltonian H, (2.1), was of the variables p_i and q_i. This invariance of form is the expression of the Galilean invariance of the underlying physical problem. If we did not have Galilean invariance, it would be impossible to find a canonical transformation which brings the transformed Hamiltonian back to its original form [except for an irrelevant added constant in (2.8)]. The formulation in terms of the once-primed coordinates and momenta is definitely not Galilean invariant, since we had to choose one preferred (moving) system of coordinates. It will be essential that a canonical transformation akin to (2.7) is *not* possible for uniform *angular* motion of the container.

We now substitute (2.7) into the definition (2.5) of P, the total linear momentum, to get

$$P=NMv+\sum_{i=1}^{N}p_{xi}''=NMv+P'', \tag{2.9}$$

where P'' is the same function of the p_i'' as P was of the p_i. P'' is clearly the linear momentum of the fluid as measured by an observer at rest with respect to the box. Substituting (2.8) and (2.9) into (2.6) we get

$$\bar{P}(v)=\frac{\exp(-\frac{1}{2}NMv^2)\,\mathrm{Trace}\{(P''+NMv)\times\exp[-\beta H(p'',q'')]\}}{\exp(-\frac{1}{2}NMv^2)\,\mathrm{Trace}\{\exp[-\beta H(p'',q'')]\}}=NMv, \tag{2.10}$$

where the last result follows from the fact that the trace of P'' is zero as a result of the formal invariance: the trace of $P''\exp[-\beta H(p'',q'')]$ is formally identical with the trace of $P\exp[-\beta H(p,q)]$, and this latter trace is just the expectation value of the total linear momentum in a stationary box, which is zero.

So far, we seem to have done no more than to prove a well-known result by a long and tedious method. However, this trivial result has nontrivial consequences. Let us evaluate the "effective mass" of the fluid in the container, i.e., the derivative of $\bar{P}(v)$ with respect to v, in the limit $v\rightarrow 0$. According to (2.10) this derivative is equal to NM. But we can also evaluate the derivative directly from Eq. (2.6). The denominator of (2.6) does not contain a term linear in v, in spite of appearances. This can be seen by observing that motion to the left and motion to the right must be equivalent in their effect on the partition function (the denominator); or else it can be seen directly from (2.8): the additional term is of order v^2. Hence we get[1]

$$\left(\frac{d\bar{P}}{dv}\right)_{v=0}=\beta\frac{\mathrm{Trace}[P^2\exp(-\beta H)]}{\mathrm{Trace}[\exp(-\beta H)]}=\beta\langle P^2\rangle_{\mathrm{Av};\ v=0}. \tag{2.11}$$

When we equate this value to the result NM obtained from (2.10), we get an equipartition theorem for the degree of freedom associated with bulk motion of the fluid:

$$\langle(P^2/2NM)\rangle_{\mathrm{Av}}=\tfrac{1}{2}kT. \tag{2.12}$$

Hence we see that the degrees of freedom associated with bulk linear motion have the equipartition value of energy, even in quantum statistical mechanics. Yet we know that the equipartition theorem itself fails in quantum statistical mechanics. What, then, is the relationship between (2.12) and the equipartition theorem in general? To see this, we take the definition of P as the sum of all the p_{xi} and substitute it into (2.12). Since the particles are all identical, all square terms give the same result, and all cross terms are equal to each other; hence we get

$$\langle P^2/2NM\rangle_{\mathrm{Av}}=\langle p_{x1}{}^2/2M\rangle_{\mathrm{Av}}+(N-1)\langle p_{x1}p_{x2}/2M\rangle_{\mathrm{Av}}=\tfrac{1}{2}kT. \tag{2.13}$$

In classical statistical mechanics, there is no correlation between momenta of different particles. That is, $\langle p_{x1}p_{x2}\rangle_{\mathrm{Av}}=0$ and the equipartition theorem is a consequence of (2.13). In quantum statistical mechanics, even an ideal gas of identical particles exhibits correlations between momenta of different particles, as a result of the Bose-Einstein (or Fermi-Dirac) statistics, i.e., as a result of the fact that the wave functions must satisfy special symmetry conditions. If we could take our trace over all wave functions, rather than over all symmetric (or antisymmetric) wave functions, the ideal (noninteracting) gas would obey the equipartition theorem even in quantum mechanics. The presence of interactions does *not* lead to correlations between momenta of different particles in classical statistical mechanics. In quantum statistical mechanics, the interparticle forces can produce momentum correlations,

---

[1] Strictly speaking, an error is made here, since the operator P does not commute with the Hamiltonian H in the presence of walls. A more detailed consideration shows that the error is of order $1/N$ (N=number of particles) and must therefore be ignored in statistical mechanics.

for example by producing bound states of pairs of particles which then form "molecules."

Equation (2.13) is the quantum-mechanical generalization of the equipartition theorem for a gas of identical particles. We can solve for the correlation function of momenta of different particles:

$$(N-1)\langle p_{x1}p_{x2}\rangle_{\text{Av}}=MkT-\langle p_{x1}^2\rangle_{\text{Av}}. \tag{2.14}$$

The deviations from classical equipartition are in opposite directions for Fermi-Dirac and Bose-Einstein statistics; the correlation coefficient $\langle p_{x1}p_{x2}\rangle_{\text{Av}}$ is negative for Fermi-Dirac statistics, positive for Bose-Einstein statistics.

A very interesting feature of (2.14) is the dependence of $\langle p_{x1}p_{x2}\rangle_{\text{Av}}$ on the number of particles N. We shall return to this point in some detail, in Sec. 5.

An equation similar to the virial theorem in classical statistical mechanics can also be derived, and again differs from the classical equation by the presence of correlation.

Finally, we can generalize to systems containing different kinds of particles. Let us take two kinds of particles, of masses M and M', respectively; there are N particles of the first kind, N' of the second kind. Then the generalization of (2.12) is

$$\langle P^2\rangle_{\text{Av}}=(NM+N'M')kT. \tag{2.15}$$

The corresponding generalization of (2.14) is not very useful, since it involves three different types of correlations.

3. DEFINITION OF THE MOMENT OF INERTIA

We now perform a similar analysis for rotational motion. Unlike the translational case, the result is not a consequence of Galilean invariance; the transformation to a rotating system of coordinates is not a Galilean transformation.

The same arguments which led us before to transform to a system of coordinates moving along with the box, now lead us to transform to a system of coordinates rotating along with the box. Just as before, the conditions of the problem *force* us to use this transformation, since otherwise we could not replace the actual walls of the container by perfectly reflecting mirror surfaces (boundary conditions on wave functions).[2]

The canonical transformation to rotating coordinates is similar to the transformation discussed in Sec. 2. We use cylindrical coordinates r, θ, z. The new coordinates are

$$\theta_i'=\theta_i-\omega t,\quad r_i'=r_i,\quad z_i'=z_i. \tag{3.1}$$

Substitution into the Lagrangean corresponding to the Hamiltonian (2.1), and transformation back to the Hamiltonian formalism gives the transformed momenta

$$p_{\theta i}'=p_{\theta i},\quad p_{ri}'=p_{ri},\quad p_{zi}'=p_{zi}. \tag{3.2}$$

The canonical variable representing the transformed angular momentum is physically the angular momentum with respect to the rest frame, *not* with respect to the moving coordinate system. The transformed Hamiltonian is

$$H'=H-\omega\sum_{i=1}^{N}p_{\theta i}=H-\omega L, \tag{3.3}$$

where L is the total angular momentum of all particles around the z-axis. Again the transformed Hamiltonian is *not* numerically equal to the original Hamiltonian, nor is it equal to the Hamiltonian as written down by an observer moving along with the box.

It is instructive to write down this latter Hamiltonian. The observer moving along with the box would first of all introduce a term in the potential energy to describe the centrifugal force, i.e., he would write down the provisional Hamiltonian

$$H_1=(2M)^{-1}\sum_{i=1}^{N}\left(p_{ri}^2+p_{zi}^2+\frac{p_{\theta i}^2}{r_i^2}\right)+V-\tfrac{1}{2}M\omega^2\sum_{i=1}^{N}r_i^2. \tag{3.4}$$

This Hamiltonian fails to give the correct equations of motion because the Coriolis force is not included. The Coriolis force is a velocity-dependent force at right angles to the velocity vector, and is therefore entirely analogous to the magnetic force on a charged particle moving in a constant magnetic field (this is of course the basis of the Larmor theorem). Hence our moving observer would introduce a "vector potential" to describe the Coriolis force; the final Hamiltonian becomes

$$H''=(2M)^{-1}\sum_{i=1}^{N}\left[p_{ri}^2+p_{zi}^2+\frac{(p_{\theta i}-Mr_i^2\omega)^2}{r_i^2}\right]+V-\tfrac{1}{2}M\omega_i^2\sum_{i=1}^{N}r_i^2. \tag{3.5}$$

It is easily seen that this Hamiltonian does give the correct equations of motion. However, H'' is formally different from the Hamiltonian of the observer in the rest system, i.e., from H, Eq. (2.1). That is, H'' is not the same function of its variables as H is of its variables. The differences are twofold: there is the additional potential energy of the centrifugal force, and the alteration in the kinetic energy associated with the Coriolis force. The centrifugal force is proportional to ω^2 and can therefore be ignored for small values of ω. One might think that we could introduce the momentum

---

[2] It is interesting to observe that, unlike the translational motion of the box, there exists a special case here in which the roughness of the walls of the box is essential to the attainment of thermal equilibrium: consider a box with cylindrical symmetry, set in rotation about its symmetry axis. If the walls of the box are perfectly reflecting mirrors, the fluid inside the box is not set into rotation at all. However, even one small irregularity on the inside surface of the box is enough to make the thermal equilibrium state unique. This small surface irregularity takes the place of the "dust particle" in the theory of blackbody radiation.

variables

$$p_{\theta i}''=p_{\theta i}-Mr_i^2\omega, \tag{3.6}$$

which correspond to the angular momentum relative to the moving system, and thereby make the Hamiltonian H'' formally identical to H, (2.1), except for terms of order ω^2 (which are of no interest for the moment of inertia). However, unlike the corresponding Galilean transformation (2.7), (3.6) is *not* a canonical transformation. The new angular momenta introduced by (3.6) fail to commute with the radial momentum variables p_{ri}, and the commutators are of order ω, not ω^2.[3] This shows the essential difference between the rotational motion and the much simpler translational motion of the box.

The thermal equilibrium value of the angular momentum of the fluid in the rotating box is given by

$$\bar{L}(\omega)=\frac{\mathrm{Trace}[L\exp(-\beta H')]}{\mathrm{Trace}[\exp(-\beta H')]}$$

$$=\frac{\mathrm{Trace}\{L\exp[-\beta(H-\omega L)]\}}{\mathrm{Trace}\{\exp[-\beta(H-\omega L)]\}}. \tag{3.7}$$

The denominator of (3.7) is the partition function of the rotating fluid, and is an even function of the angular velocity ω, by symmetry. We are interested in the moment of inertia, which is defined by

$$I=\left(\frac{d\bar{L}}{d\omega}\right)_{\omega=0}. \tag{3.8}$$

Unlike the translational case, where $\bar{P}$ is a linear function of v as long as v is small compared to the velocity of light, the angular momentum $\bar{L}$ is not a purely linear function of ω; as the angular velocity increases, centrifugal forces come into play, particles tend to be forced outward, and $\bar{L}$ increases more rapidly; that is, the effective moment of inertia is itself a function of ω. However, these effects are not involved in the definition (3.8) since we go to the limit $\omega\rightarrow 0$.

We now differentiate (3.7) with respect to ω, and then set ω equal to 0. The denominator is an even function of ω, and therefore its derivative with respect to ω is zero at $\omega=0$. We then get the following definition of the moment of inertia of a fluid in thermal equilibrium:

$$I=\beta\frac{\mathrm{Trace}[L^2\exp(-\beta H)]}{\mathrm{Trace}[\exp(-\beta H)]}, \tag{3.9}$$

where H, Eq. (2.1), is the Hamiltonian of the fluid in a *stationary* container.

The following sections contain a proof, in classical and then quantum statistical mechanics, that I defined by (3.9) is equal to the conventional expression for the moment of inertia, Eq. (1.2).

4. THE MOMENT OF INERTIA IN CLASSICAL STATISTICAL MECHANICS

We first prove a lemma which is itself of some interest since it represents a considerable generalization of the equipartition theorem. To save writing, we introduce coordinates $q_1=x_1$, $q_2=y_1$, $\cdots$, $q_{3N}=z_N$, and we label the corresponding momentum components accordingly. Then the following lemma holds:

Lemma: Let $f(q_1,\cdots,q_{3N})$ be any function of the coordinates only. Then

$$\langle p_kp_{k'}f\rangle_{\mathrm{Av}}=\delta_{kk'}M_kkT\langle f\rangle_{\mathrm{Av}}. \tag{4.1}$$

The proof of the lemma is patterned along the conventional proof of the equipartition theorem. The average value on the left side of (4.1) is given by

$$\langle p_kp_{k'}f\rangle_{\mathrm{Av}}=\frac{\int d^{3N}q\int d^{3N}p\; p_kp_{k'}f\exp(-\beta H)}{\int d^{3N}q\int d^{3N}p\exp(-\beta H)}. \tag{4.2}$$

The Hamiltonian is a diagonal quadratic form in the momentum components, i.e.,

$$H=\sum_{k=1}^{3N}\frac{p_k^2}{2M_k}+V(q_1,\cdots,q_{3N}). \tag{4.3}$$

Hence we have the identity

$$p_k=M_k\partial H/\partial p_k. \tag{4.4}$$

Substitution into the numerator of (4.2) gives

$$\text{Numerator}=-(M_k/\beta)\int d^{3N}q\int d^{3N}p\; fp_{k'}\frac{\partial\exp(-\beta H)}{\partial p_k}$$

$$=+\delta_{kk'}(M_k/\beta)\int d^{3N}q\int d^{3N}p\; f\exp(-\beta H), \tag{4.5}$$

where the second step comes from an integration by parts on p_k together with the observation that the function f is independent of p_k. Combination of (4.5) and (4.2) proves the lemma.

We are now in a position to prove the main theorem. We need the average value of L^2, and we therefore write

$$L^2=\left(\sum_{i=1}^{N}(x_ip_{yi}-y_ip_{xi})\right)^2$$

$$=\sum_{i,j=1}^{N}(x_ip_{yi}-y_ip_{xi})(x_jp_{yj}-y_jp_{xj}). \tag{4.6}$$

---

[3] An expansion in powers of this commutator is possible, however, as has been shown by R. Peierls, Z. Physik **80**, 763 (1933).

Each term of (4.6) is of the form appropriate for use of our lemma (4.1). When we employ the lemma, all nondiagonal terms drop out identically, and we get

$$I=\beta\langle L^2\rangle_{\rm Av}=\sum_{i=1}^{N} M_i\langle(x_i^2+y_i^2)\rangle_{\rm Av} \tag{4.7}$$

which is what we set out to prove.

At first sight it may seem as if we have proved too much; for consider the following counter example: a cylindrical, closed container is filled partially with water and is mounted with its axis horizontal. The water settles in the lower part of the container. We now start rotating the container about its horizontal axis. The water stays near the bottom, and the observed angular momentum is not equal to $I\omega$ with I given by (4.7). The answer is that under these conditions thermal equilibrium is never reached. As the container rotates, the layer of water in contact with the wall moves along with the wall, and closed currents are set up in the water. These currents maintain themselves as long as the container is kept rotating. The viscous forces then produce heat in the water, and the temperature increases. Thus our theorem is not in contradiction to this case, simply because the theorem starts from the assumption of thermal equilibrium and is therefore not applicable to this case.

One condition of applicability of our theorem, i.e., the condition for the existence of a rotating motion of the fluid in thermal equilibrium, is that the potential energy V must be invariant under rotations.[4]

This is true, for example, for a rotating bucket of water with the axis of rotation parallel to the direction of the gravitational force, and is not true for a bucket of water mounted with the axis horizontal.

5. CORRELATION LENGTH

Unlike classical statistical mechanics, Eq. (4.7) does not hold exactly in quantum statistical mechanics. For normal systems, we shall show that (4.7) is approximately correct, the approximation becoming better and better as the size of the container is increased. However, there exists at least one exceptional system (the ideal Bose-Einstein gas below its condensation point) in which (4.7) is violated, no matter how big the size of the container.[5] It is clear, therefore, that some physical assumption about the fluid is necessary in order to prove (4.7) in quantum statistical mechanics. Thermal equilibrium alone is not enough.

We shall restrict ourselves to fluids (gases and liquids), and ignore solids. The moment of inertia of a solid is equal to the classical value immediately, as a result of the solidity, so we do not lose much in generality by restricting ourselves to fluids from the start. We point out, however, that our considerations *can* be applied to the electron gas in a metal, thereby enabling us to deduce some statements of interest for the theory of superconductivity (this will form the subject of a separate paper).

Let us consider two particles in our field, separated by a distance r_{12}. The correlation coefficient between their momentum components p_{x1} and p_{x2}, say, is defined by

$$Q_{xx}(r_{12})=\langle p_{x1}p_{x2}\rangle_{\rm Av}/\langle(p_{x1}^2)\rangle_{\rm Av}^{\frac{1}{2}}\langle(p_{x2}^2)\rangle_{\rm Av}^{\frac{1}{2}}$$

with similar equations for Q_{xy}, etc. We assert that these correlation functions depend on the distance r_{12} between the two particles in the fluid, and in particular there must exist in any physical fluid a "correlation distance" Λ such that the correlations Q become negligibly small for $r_{12}\gg\Lambda$. Since this correlation distance forms the starting point of the entire subsequent discussion, we shall now devote some time to clarifying its meaning.

First of all, the correlations Q are identically zero in classical statistical mechanics. That is, two different particles of the fluid always have uncorrelated momentum components in thermal equilibrium, no matter how close together the particles are. Thus the correlation length Λ in classical statistical mechanics is zero. One might think *a priori* that momentum components of different particles should show a correlation if the particles are within one mean free path of each other, but the formalism of classical statistical mechanics shows that this is not so. The mean free path is a nonequilibrium concept, whereas the correlation length is an equilibrium quantity, and in classical statistical mechanics there is no relation between the two.

In quantum statistical mechanics, one may at first object to our definition of the correlation coefficients, on the grounds that the momenta and coordinates of the particles don't commute. However, it is still possible to define a quantity analogous to a probability of finding one particle at $\mathbf{r}_1$ with momentum $\hbar\mathbf{k}_1$ and another particle at $\mathbf{r}_2$ with momentum $\hbar\mathbf{k}_2$. This is done in the next section, Eq. (6.20). The analogy is imperfect, since the quantum mechanical expression is not only not positive definite, but even fails to be real. However, this does not affect the possibility of defining a correlation length, since the correlation coefficient Q can become small without being real or positive.

It was pointed out in Sec. 2 that there are two ways in which quantum mechanics leads to momentum correlations between particles of the fluid. First, there is the possibility of the formation of discrete bound states (molecules) in which the momenta of the atoms making up one molecule are strongly correlated. Second, the requirements of Bose-Einstein or Fermi-Dirac statistics entail momentum correlations even in an ideal

---

[4] There are two kinds of "rotational invariance": (1) Invariance of the form of the Hamiltonian under change of the origin of the angle variable θ, i.e., under the transformation $\theta'=\theta-\alpha$, and (2) Invariance of the form of the Hamiltonian under a transformation to rotating coordinates, $\theta'=\theta-\omega t$. The first invariance is the one we need here, and is frequently fulfilled in practice. The second invariance is never true. This differs from linear motion, where both translational invariance and Galilean invariance are observed.

[5] J. M. Blatt and S. T. Butler, preceding paper [Phys. Rev. **100**, 476 (1955)].

gas. Let us make a very rough first estimate of the correlation lengths we can expect from these two effects: the first effect gives rise to momentum correlations over distances of the order of magnitude of interatomic distances within molecules (we exclude solids from this discussion), the second effect can be estimated to give momentum correlations of significance over distances of the order of a de Broglie wavelength of a particle with energy $E=kT$.

These rough estimates are probably adequate for most systems, but they fail for some special systems. The most conspicuous failure occurs for the ideal Bose-Einstein gas below its condensation point. This system contains a macroscopic number of particles (comparable to the total number N) in a single-particle state of de Broglie wavelength comparable to the size of the box. Thus this de Broglie wavelength, rather than the de Broglie wavelength of a particle with energy $E=kT$, gives the most far-reaching momentum correlations, and momenta of different particles in the gas are correlated even when the particles are at opposite ends of the box, no matter how big we make the box. That is, the ideal Bose-Einstein gas below its condensation point does not have a finite correlation length (a correlation length independent of the size of the container).

The other ideal gases (Bose-Einstein gas above the condensation point, Fermi-Dirac gas) have in principle the possibility of infinite-range momentum correlations, because the same quantum state which is responsible for the long-range correlation effects in the condensed Bose gas exists in these other gases and can be occupied by particles of the gas. However, in the other ideal gases this quantum state, as well as all the other very long wavelength states, has such a small occupation probability that no significant correlation results from it.

When we turn from ideal gases (no interactions between the particles) to real physical fluids, one significant difference appears immediately: the interactions between the particles make correlations of *infinite* range (range as large as the container) impossible. For suppose the opposite were true; we could then take a volume of fluid as large as the interior of the sun, and would have to expect to find individual atoms with appreciably correlated momenta at opposite ends of this volume, in thermal equilibrium; if this be admitted as a possibility, let us make the container as large as the interior of a red giant star, or as large as the whole galaxy. It is clear that, no matter how weak the interactions between the particles are, they must eventually wash out correlation effects between very distant atoms of the fluid.

As a rough first approximation, applicable to dilute gases, we may assume that an upper limit to the correlation distance Λ is given by a few times the mean free path of a particle. That is, we assume that two particles which are many mean free paths away from each other have statistically uncorrelated momenta.[6] Of course, this upper limit may be much too high (as is shown by the example of classical statistical mechanics, where $\Lambda=0$ no matter how large the mean free path) and is not applicable to liquids, since the mean free path is not a useful concept in the liquid state. An upper limit to Λ for liquid helium will be established by special arguments in a subsequent paper. For the moment, we need not estimate the value of Λ; all we need is that a correlation distance Λ *exists* (i.e., is independent of the volume of the container) for any physical fluid.

There is then a significant qualitative difference between real fluids, no matter how weakly the particles interact with each other, and ideal, noninteracting gases. Real fluids always possess a correlation length Λ, and two particles in a real fluid are statistically uncorrelated when they are apart by a distance $r \gg \Lambda$; ideal gases *may* show long range correlations, and the ideal Bose-Einstein gas below its condensation point actually does so.[7]

It should be pointed out that the correlation length Λ is a quantity defined in the thermodynamic equilibrium state, not a transfer property. It is of course well known that the mean free path is of great importance in transfer phenomena; for example the flow of the fluid through a pipe of dimensions large compared to the mean free path is qualitatively different from the flow through a narrow pipe of diameter less than the mean free path (Knudsen regime). We are *not* concerned with this, but are pointing out qualitative differences between a "Knudsen gas" and a normal gas *in thermodynamic equilibrium.*

The ideal Bose-Einstein gas below its condensation point has a nonclassical moment of inertia[5]; we know from experience that most ordinary fluids have the classical moment of inertia. The discussion above has shown that there exists another property of fluids, the correlation length, which distinguishes the ideal Bose-Einstein gas below condensation from all other fluids. It is reasonable, therefore, to suppose that there is a connection between the moment of inertia and the correlation length. We shall show in the subsequent

---

[6] A single mean free path may not be enough, since significant momentum correlations might be maintained as the result of small chains of collisions. In principle, the correlation coefficient never becomes identically zero; however, the residual correlation coefficient at large distances decreases very rapidly with the distance (presumably exponentially) and therefore becomes negligible very quickly.

[7] The dangers of ignoring the infinite correlation length can be seen for example in the paper of P. R. Zilsel, Phys. Rev. **92**, 1106 (1953). Zilsel attempts to impose an external condition (prescribed value of the total linear momentum P) on a *part* of the volume of a container filled with an ideal Bose-Einstein gas below its condensation point. He then obtains results inconsistent with Galilean invariance (see Sec. 2 of this paper). The existence of a correlation length Λ in a real Bose-Einstein fluid implies important qualitative changes in the nature of the Bose-Einstein condensation phenomenon itself. However these qualitative changes are not of experimental importance provided the correlation length Λ is much larger than the mean distance between particles in the fluid.

mathematical sections that this is indeed the case. To be precise, let I be the actual moment of inertia defined by (3.9), and let I_0 be the classical moment of inertia, $\frac{1}{2}NMR^2$ for a cylindrical bucket of radius R. We shall show that[8]

$$I=I_0[1+\text{order}(\Lambda S/V)], \tag{5.1}$$

where S is the surface area of the container and V is its volume. If the fluid is normal, i.e., possesses a definite correlation length Λ, the correction term can be made arbitrarily small simply by increasing all linear dimensions of the container. But in the ideal Bose-Einstein gas below its condensation point, Λ is of the same order as the linear dimensions of the container, and the correction term remains important no matter how big we make the container.

In the later application, we shall need a somewhat stronger form of (5.1) for the special case that the container has the form of a cylinder of radius R and height D. We shall show in Appendix A that for such a cylinder the height D is unimportant, and hence

$$I=I_0[1+\text{order}(\Lambda/R)]. \tag{5.2}$$

The proof is relegated to an appendix because the formalism is simpler if *all* linear dimensions of the container, including the height, are large compared to the correlation length.

Having described our program, we continue the discussion of the correlation length. As soon as the particles in the fluid interact with each other, it is impossible to write the wave functions of the actual fluid as products of single-particle wave functions. Thus we can no longer talk of "occupation numbers" of single-particle states, and the free energy of the fluid cannot be written as a sum of contributions from single-particle states. However, it is still possible to define a quantity analogous to a probability of finding a particle at the point $\mathbf{r}$ with momentum equal to $\hbar\mathbf{k}$.[9]

It will be important in our subsequent mathematical work to use plane waves as our basic set of functions. Since the plane wave functions fail to satisfy the correct boundary conditions at the walls of the container, an error is made by this procedure, and this error is of the order of $\lambda_m S/V$, where λ_m is the largest de Broglie wavelength of importance in the partition function, S is the surface area of the container, and V is the volume of the container. This error estimate is the ratio of the volume in which the error is made, to the total volume. We shall now give arguments to prove that λ_m, the maximum de-Broglie wavelength of importance, is less than the correlation distance Λ.

We have mentioned two sources of momentum correlations in quantum statistical mechanics. One of these was the requirement of symmetry or antisymmetry of the wave functions. This requirement imposes momentum correlations of particles within one de Broglie wavelength of each other. Thus the correlation distance Λ must be at least as large as the maximum de Broglie wavelength of importance, i.e., $\lambda_m \leqslant \Lambda$. Of course, the correlation distance Λ may be much larger than this in special systems, but this does not concern us here.

Finally, we return to formula (2.14); since the average $\langle p_{x1}{}^2\rangle_{\text{Av}}$ is an intensive quantity the correlation coefficient $\langle p_{x1}p_{x2}\rangle_{\text{Av}}$ is seen to be inversely proportional to the number of particles in the system. In a fluid with correlation length Λ, this dependence is easily understood: when the two particles are farther apart than Λ, their momenta are uncorrelated, and the contribution to $\langle p_{x1}p_{x2}\rangle_{\text{Av}}$ vanishes. Thus the only regions of the two-particle configuration space $(\mathbf{r}_1,\mathbf{r}_2)$ which contribute significantly to $\langle p_{x1}p_{x2}\rangle_{\text{Av}}$ are such that $|\mathbf{r}_1-\mathbf{r}_2|<\Lambda$. The contributing volume is therefore $\Lambda^3 V$ rather than V^2, and the correlation coefficient $\langle p_{x1}p_{x2}\rangle_{\text{Av}}$ is proportional to Λ^3/V, i.e., inversely proportional to the volume and hence also inversely proportional to the number of particles N, in agreement with (2.14).

While (2.14) can be understood in terms of the correlation length, we emphasize that (2.14) holds even when no correlation length can be defined, for example, (2.14) holds for the ideal Bose-Einstein gas below its condensation point. The reason is that correlations, even over large distances, between *very small* momenta p_1 and p_2 make no effective contribution either to $\langle p_1{}^2\rangle_{\text{Av}}$ or to $\langle p_1p_2\rangle_{\text{Av}}$. Since long-range correlations are important only for states with long de Broglie wavelengths, and hence with small momenta, (2.14) does not depend for its validity upon the existence of a finite correlation length, and conversely the validity of (2.14) cannot be used to prove the existence of a finite correlation length for any particular fluid.

6. AVERAGES OF ONE-PARTICLE AND TWO-PARTICLE OPERATORS

In this section, we develop some general methods necessary to carry through the proof of (4.7) in quantum

---

[8] This theorem is correct, and will be proved to be so, on the assumption that the particles have no intrinsic spin (which is true for helium atoms). If the particles do have intrinsic spin, this spin may contribute to the total angular momentum of the rotating fluid, as a result of spin-orbit coupling. A more detailed consideration shows that this contribution is of order ω^3 for all systems except perhaps ferromagnets; a contribution of order ω^3 to the angular momentum is of no importance for the moment of inertia, Eq. (3.8). More important, however, is the fact that the angular momentum contributed by the spins is at most $Ns\hbar$, where N is the number of particles. (This value is obtained when all spins are lined up parallel to each other.) As the radius of the bucket is increased, the orbital contribution to the angular momentum is proportional to NR^2, hence for large enough buckets the spin contribution is always negligible compared to the orbital contribution. Since the spins are of no importance for the final result, we shall ignore spins from now on.

[9] This is the quantity $W(\mathbf{k}_1,\mathbf{r}_1)$ defined in Sec. 6. The analogy is not perfect because W can assume negative values. However, the expectation value of any operator F which is a sum of single-particle operators, $F=\sum_{i=1}^{N}F_i$, can be expressed in terms of W.

statistical mechanics. These methods are not without some interest apart from this particular problem.[10]

Let F be a sum of single-particle operators

$$F=\sum_{i=1}^{N} F_i, \tag{6.1}$$

where each F_i acts only on the coordinates of the ith particle. We are interested in the expectation value of F in thermal equilibrium:

$$\bar{F}=\frac{\text{Trace}[F\exp(-\beta H)]}{\text{Trace}[\exp(-\beta H)]}=N\frac{\text{Trace}[F_1\exp(-\beta H)]}{\text{Trace}[\exp(-\beta H)]}, \tag{6.2}$$

where the trace is to be taken over all symmetric wave functions for Bose-Einstein statistics, over all antisymmetric wave functions for Fermi-Dirac statistics.

We shall attempt to simplify the evaluation of this trace under the *assumption that there exists a correlation distance* Λ, of the kind discussed in Sec. 5. Let S be the surface area of the volume V in which our system is enclosed. We define the function $G(\mathbf{k},\mathbf{r})$ by

$$G(\mathbf{k},\mathbf{r})=\tfrac{1}{2}\exp(-i\mathbf{k}\cdot\mathbf{r})F_1\exp(+i\mathbf{k}\cdot\mathbf{r}) +\text{complex conj.} \tag{6.3}$$

We assert that there exists a real function $W(k^2)$ such that

$$\text{Trace}[F_1\exp(-\beta H)]=\sum_{\mathbf{k}_1}\int\frac{d^3r_1}{V}G(\mathbf{k}_1,\mathbf{r}_1)W(k_1{}^2) +\text{terms of order } \Lambda S/V. \tag{6.4}$$

The sum over $\mathbf{k}_1$ is a sum over plane wave states in the volume V; it can often be replaced by an integral over $\mathbf{k}_1$ space. The function W is the same for all operators F, and it is therefore in some sense a probability density for finding the wave vector $\mathbf{k}$ in an interval d^3k. However, W is not necessarily positive.

The proof of (6.4) proceeds in several stages. First of all, we have shown in Sec. 5 that the existence of a correlation distance allows us to perform the trace by summing over all plane wave functions.

Next we introduce notation. We let $\mathbf{R}$ stand for the ordered set of coordinate vectors

$$\mathbf{R}=(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N), \tag{6.5}$$

and $\mathbf{K}$ for the ordered set of wave vectors

$$\mathbf{K}=(\mathbf{k}_1,\mathbf{k}_2,\cdots,\mathbf{k}_N), \tag{6.6}$$

and we use the short-hand notation

$$\exp(i\mathbf{K}\cdot\mathbf{R})=\exp[i(\mathbf{k}_1\cdot\mathbf{r}_1+\mathbf{k}_2\cdot\mathbf{r}_2+\cdots+\mathbf{k}_N\cdot\mathbf{r}_N)]. \tag{6.7}$$

Let Π be the permutation which replaces 1 by s_1, 2 by s_2, $\cdots$, N by s_N. Then the vector $\mathbf{R}^{\Pi}$ is defined by

$$\mathbf{R}^{\Pi}=(\mathbf{r}_{s1},\mathbf{r}_{s2},\cdots,\mathbf{r}_{sN}). \tag{6.8}$$

The permutation operator Π applied to the plane wave $\exp(i\mathbf{K}\cdot\mathbf{R})$ gives

$$\Pi\exp(i\mathbf{K}\cdot\mathbf{R})=\exp(i\mathbf{K}\cdot\mathbf{R}^{\Pi}). \tag{6.9}$$

We define ϵ_{Π} to equal $+1$ for Bose-Einstein statistics, to equal $(-1)^{\Pi}$ for Fermi-Dirac statistics. Furthermore, we define the symbol TRACE to mean a trace over all states, irrespective of symmetry; the symbol Trace refers to a trace over symmetric (Bose-Einstein statistics) or antisymmetric (Fermi-Dirac statistics) states only. We then get[11]

$$\begin{aligned}\text{Trace}[F_1\exp(-\beta H)]&=(N!)^{-1}\sum_{\Pi}\epsilon_{\Pi}\,\text{TRACE}[F_1\exp(-\beta H)\Pi]\\&=\tfrac{1}{2}(N!)^{-1}\sum_{\Pi}\epsilon_{\Pi}\,\text{TRACE}[F_1\exp(-\beta H)\Pi\\&\qquad+\Pi\exp(-\beta H)F_1],\end{aligned} \tag{6.10}$$

where the sum extends over all possible permutations Π, and the symmetrized form is preferable to exhibit manifest reality of the final quantities.

We now write down the trace in completely explicit form. There is a sum over all plane wave states, i.e., over all sets of vectors $\mathbf{k}_1$, $\mathbf{k}_2$, $\cdots$, $\mathbf{k}_N$. Furthermore, there is an integration over $3N$ dimensional configuration space to obtain the diagonal matrix elements for the trace. Thus we get, using the Hermitean properties of F_1, H, and Π

$$\begin{aligned}\text{Trace}[F_1\exp(-\beta H)]&=\tfrac{1}{2}(N!)^{-1}\sum_{\Pi}\epsilon_{\Pi}\sum_{\mathbf{K}}\int\frac{d^{3N}r}{V^N}\\&\times\{[F_1\exp(i\mathbf{K}\cdot\mathbf{R})]^*\exp(-\beta H)\exp(i\mathbf{K}\cdot\mathbf{R}^{\Pi})\\&+[\exp(-\beta H)\exp(i\mathbf{K}\cdot\mathbf{R}^{\Pi})]^*F_1\exp(i\mathbf{K}\cdot\mathbf{R})\}.\end{aligned} \tag{6.11}$$

We define the function $G_1(\mathbf{k}_1,\mathbf{r}_1)$ by the relation

$$\begin{aligned}G_1(\mathbf{k}_1,\mathbf{r}_1)&=\exp(-i\mathbf{K}\cdot\mathbf{R})F_1\exp(i\mathbf{K}\cdot\mathbf{R})\\&=\exp(-i\mathbf{k}_1\cdot\mathbf{r}_1)F_1\exp(i\mathbf{k}_1\cdot\mathbf{r}_1).\end{aligned} \tag{6.12}$$

The second form follows because the operator F_1 acts upon the coordinates of particle 1 only, by assumption. We then get

$$\begin{aligned}\text{Trace}[F_1\exp(-\beta H)]&=\tfrac{1}{2}(N!)^{-1}\sum_{\Pi}\epsilon_{\Pi}\sum_{\mathbf{K}}\int\frac{d^{3N}r}{V^N}\\&\times[G_1{}^*(\mathbf{k}_1,\mathbf{r}_1)\exp(-i\mathbf{K}\cdot\mathbf{R})\exp(-\beta H)\\&\times\exp(i\mathbf{K}\cdot\mathbf{R}^{\Pi})+\text{compl. conj.}].\end{aligned} \tag{6.13}$$

---

[10] Similar methods have been used elsewhere: E. P. Wigner, Phys. Rev. **44**, 31 (1931); J. E. Moyal, Proc. Cambridge Phil. Soc. **45**, 99 (1949); J. E. Mayer and W. Band, J. Chem. Phys. **15**, 141 (1947); H. S. Green, Proc. Roy. Soc. (London) **A194**, 244 (1948); J. Chem. Phys. **19**, 955 (1951). Our main interest here is in the symmetry properties of the functions W and U defined in this section, rather than in an explicit evaluation of these functions.

[11] B. Kahn and G. E. Uhlenbeck, Physica **5**, 399 (1938).

We now observe that the integrations over $\mathbf{r}_2$, $\mathbf{r}_3$, $\cdots$, $\mathbf{r}_N$ do not involve the function G_1 at all. The same holds true of the sums over $\mathbf{k}_2$, $\mathbf{k}_3$, $\cdots$, $\mathbf{k}_N$, and of the sum over permutations Π. Thus all these operations can be performed once and for all, independently of the form of G_1, i.e., independently of the operator F_1. We define a function $W(\mathbf{k}_1,\mathbf{r}_1)$ by

$$W(\mathbf{k}_1,\mathbf{r}_1)=(N!)^{-1}\sum_{\Pi}\epsilon_{\Pi}\sum_{\mathbf{k}_2\cdots\mathbf{k}_N}\int\frac{d^3r_2\cdots d^3r_N}{V^{N-1}} \times\exp(-i\mathbf{K}\cdot\mathbf{R})\exp(-\beta H)\exp(+i\mathbf{K}\cdot\mathbf{R}^{\Pi}). \quad (6.14)$$

This is a definite function of $\mathbf{k}_1$ and $\mathbf{r}_1$, *not* a formal quantity which still involves operators acting on r_1. For $\exp(-\beta H)\exp(i\mathbf{K}\cdot\mathbf{R}^{\Pi})$ is a definite function of the variables $\mathbf{r}_1$, $\mathbf{r}_2$, $\cdots$, $\mathbf{r}_N$, and this is multiplied by $\exp(-i\mathbf{K}\cdot\mathbf{R})$ and integrated over $\mathbf{r}_2$, $\cdots$, $\mathbf{r}_N$, thereby yielding a definite function of $\mathbf{r}_1$. $W(\mathbf{k}_1,\mathbf{r}_1)$ is very hard to evaluate explicitly; but we shall not need to do so in this paper.

We also observe that we have *not* replaced sums over wave vectors $\mathbf{k}$ by integrals over a continuous k-space. It is well known that this procedure leads to difficulties when the particles in the system can form bound states, and results in an incorrect omission of the contributions of the bound states to the trace. But as long as we keep to sums over discrete wave vectors $\mathbf{k}$, we are summing over a complete set of functions, and no states are omitted.

With this definition, we get

$$\mathrm{Trace}[F_1\exp(-\beta H)]=\tfrac{1}{2}\sum_{\mathbf{k}_1}\int\frac{d^3r_1}{V}[G_1^*(\mathbf{k}_1,\mathbf{r}_1)W(\mathbf{k}_1,\mathbf{r}_1) +W^*(\mathbf{k}_1,\mathbf{r}_1)G_1(\mathbf{k}_1\mathbf{r}_1)]. \quad (6.15)$$

We now proceed to show that, to the approximation involved in neglecting terms of order $\Lambda S/V$, the function $W(\mathbf{k}_1,\mathbf{r}_1)$ is independent of $\mathbf{r}_1$, depends on $\mathbf{k}_1$ only through k_1^2, and is real.

To show this, we observe that the existence of the correlation distance Λ allows us, in the calculation of W, (6.14), to replace the true volume first of all by a cube of the same total volume, and then to replace the correct boundary conditions at the surfaces of the cube by the condition of periodicity of all wave functions. Both operations introduce errors of order $\Lambda S/V$, and are therefore allowed to us.

We now show that $W(\mathbf{k}_1,\mathbf{r}_1)$ is independent of $\mathbf{r}_1$ in the periodicity cube, provided only that the Hamiltonian operator H is invariant under all translations.[12] Let us consider the function $W(\mathbf{k}_1, \mathbf{r}_1+\mathbf{a})$ where $\mathbf{a}$ is a constant vector. In the definition (6.14) of W, we introduce new variables of integration: $\mathbf{r}_2'=\mathbf{r}_2-\mathbf{a}$, $\mathbf{r}_3'=\mathbf{r}_3-\mathbf{a}$, $\cdots$ $\mathbf{r}_N'=\mathbf{r}_N-\mathbf{a}$. Since we have imposed periodic boundary conditions, we do not need to change the limits of integration. The function $\exp(-i\mathbf{K}\cdot\mathbf{R})$ is then changed into

$$\exp\left[-i\sum_{j=1}^{N}\mathbf{k}_j\cdot\mathbf{a}\right]\exp(-i\mathbf{K}\cdot\mathbf{R}),$$

and the function $\exp(+i\mathbf{K}\cdot\mathbf{R}^{\Pi})$ is changed into

$$\exp\left[+i\sum_{j=1}^{N}\mathbf{k}_j\cdot\mathbf{a}\right]\exp(+i\mathbf{K}\cdot\mathbf{R}^{\Pi}).$$

The additional factors are just constants as far as the integrations over $\mathbf{r}_2$, $\mathbf{r}_3$, $\cdots$, $\mathbf{r}_N$ are concerned, and also as far as the operator $\exp(-\beta H)$ is concerned. Thus these factors cancel, and we conclude that $W(\mathbf{k}_1,\mathbf{r}_1)$ is independent of $\mathbf{r}_1$. Furthermore, W is a scalar function of the vector $\mathbf{k}_1$ only (it is independent of $\mathbf{r}_1$), and this means it must be a function of k_1^2 only.[13]

Finally, we observe that an additional integration of (6.14) over $\mathbf{r}_1$ makes the integral in (6.14) into an integration over the full $3N$-dimensional configuration space. We conclude from the Hermitian property of the operator $\exp(-\beta H)$ and of the operator $\sum_{\Pi}\epsilon_{\Pi}\Pi$ that

$$\int d^3x_1W(\mathbf{k}_1,\mathbf{r}_1)=\int d^3x_1W^*(\mathbf{k}_1,\mathbf{r}_1). \quad (6.16)$$

Since W is independent of r_1, we conclude that W is a real function of k_1^2. This finishes the proof of the statements made in connection with Eq. (6.4).

We now turn to expectation values of operators involving the coordinates of *two* particles. Let F be an operator of type

$$F=\sum_{\substack{i,j=1\\(i\neq j)}}^{N}F_{ij}, \quad (6.17)$$

where F_{ij} involves the coordinates of particles i and j only. We have

$$\mathrm{Trace}[F\exp(-\beta H)] =N(N-1)\,\mathrm{Trace}[F_{12}\exp(-\beta H)]. \quad (6.18)$$

---

[12] For our later use of these theorems, this condition is stronger than necessary, and is not always satisfied. For example, the Hamiltonian of water in a bucket under the action of gravity is not invariant under translations parallel to the force of gravity. We shall need only invariance under translations perpendicular to the axis of rotation of the bucket, and hence can allow $W(\mathbf{k}_1,\mathbf{r}_1)$ to depend upon z_1, but not upon x_1 or y_1. The modifications of the proof introduced by the possibility of these weaker conditions are rather trivial, and we shall indicate them in the Appendix.

[13] The reader may object on one of two grounds: (a) this is so trivial it needs no proof; (b) the proof is incorrect since the first step, i.e., the replacement of the actual volume by a cube of the same volume content, already assumed that W was independent of r_1, i.e., we have assumed at the start what we tried to prove. Objection (a) can be countered by observing that it is not obvious *a priori* that $W(\mathbf{k}_1,\mathbf{r}_1)$ might not be complex with a phase factor $\exp(i\mathbf{k}_1\cdot\mathbf{r}_1)$; this has been excluded by our proof. Objection (b) is incorrect: the correlation distance Λ assures us that $W(\mathbf{k}_1,\mathbf{r}_1)$ is a "local" property, i.e., W depends only on conditions within a neighborhood of size Λ around the point $\mathbf{r}_1$; but a local property need not be independent of the locality. The replacement of the actual volume by a cube is possible because W is a local property, and does *not* presume that W is independent of position.

We define the function $G(\mathbf{k}_1,\mathbf{k}_2,\mathbf{r}_1,\mathbf{r}_2)$ by an equation analogous to (6.12):

$$\begin{aligned}G(\mathbf{k}_1,\mathbf{k}_2,\mathbf{r}_1,\mathbf{r}_2)&=\exp(-i\mathbf{K}\cdot\mathbf{R})F_{12}\exp(i\mathbf{K}\cdot\mathbf{R})\\&=\exp[-i(\mathbf{k}_1\cdot\mathbf{r}_1+\mathbf{k}_2\cdot\mathbf{r}_2)]F_{12}\\&\quad\times\exp[+i(\mathbf{k}_1\cdot\mathbf{r}_1+\mathbf{k}_2\cdot\mathbf{r}_2)],\end{aligned}\tag{6.19}$$

and a function $U(\mathbf{k}_1,\mathbf{k}_2,\mathbf{r}_1,\mathbf{r}_2)$ by an equation analogous to (6.14):

$$\begin{aligned}U(\mathbf{k}_1,\mathbf{k}_2,\mathbf{r}_1,\mathbf{r}_2)&\\=(N!)^{-1}\sum_{\Pi}\epsilon_{\Pi}\sum_{\mathbf{k}_3,\mathbf{k}_4,\cdots,\mathbf{k}_N}&\int\frac{d^3r_3d^3r_4\cdots d^3r_N}{V^{N-2}}\\\times\exp(-i\mathbf{K}\cdot\mathbf{R})\exp(-\beta H)&\exp(+i\mathbf{K}\cdot\mathbf{R}^{\Pi}).\end{aligned}\tag{6.20}$$

Comparison of (6.14) and (6.20) gives the relation

$$W(\mathbf{k}_1,\mathbf{r}_1)=\sum_{\mathbf{k}_2}\int\frac{d^3x_2}{V}U(\mathbf{k}_1,\mathbf{k}_2,\mathbf{r}_1,\mathbf{r}_2).\tag{6.21}$$

By writing down the trace in (6.18) explicitly, we see that

$$\mathrm{Trace}[F_{12}\exp(-\beta H)]=\tfrac{1}{2}\sum_{\mathbf{k}_1,\mathbf{k}_2}\int\frac{d^3r_1d^3r_2}{V^2}(G^*U+U^*G).\tag{6.22}$$

We now establish certain properties of the function U which we shall need in the next section. First of all, the definition of U, (6.20), together with the assumption of a correlation distance Λ independent of the volume of the container, allows us to assert that U is invariant under translations of $\mathbf{r}_1$ and $\mathbf{r}_2$ simultaneously. The detailed argument is exactly the same as for W and needs no repetition. Next, by our fundamental assumption of a correlation distance Λ, U becomes separable if the distance $|\mathbf{r}_1-\mathbf{r}_2|$ exceeds Λ. That is

$$U(\mathbf{k}_1,\mathbf{k}_2,\mathbf{r}_1,\mathbf{r}_2)=W(k_1^2)W(k_2^2)\quad\text{for}\quad|\mathbf{r}_1-\mathbf{r}_2|\gg\Lambda.\tag{6.23}$$

The fact that the factors are just $W(k_1^2)$ follows from separability together with Eq. (6.21). From the translation symmetry we conclude that

$$U(\mathbf{k}_1,\mathbf{k}_2,\mathbf{r}_1,\mathbf{r}_2)=U(\mathbf{k}_1,\mathbf{k}_2,\mathbf{r}_{12})\quad[\text{except near the boundaries}].\tag{6.24}$$

When $\mathbf{r}_1$ and $\mathbf{r}_2$ are well separated, (6.23) applies and U is real since W is real. When $|\mathbf{r}_{12}|$ is of order Λ or smaller, U need not be real. However, U has additional symmetries as a result of being a "local" property: U is invariant under rotations of the set of vectors $\mathbf{k}_1$, $\mathbf{k}_2$, $\mathbf{r}_{12}$ simultaneously, and it is also invariant under the inversion operation:

$$U(-\mathbf{k}_1,-\mathbf{k}_2,-\mathbf{r}_{12})=U(\mathbf{k}_1,\mathbf{k}_2,\mathbf{r}_{12}).\tag{6.25}$$

This finishes our general analysis of expectation values of one- and two-particle operators. A word of caution is perhaps in order. The theorems may look more impressive than they really are. For most practical problems we need to know the temperature dependence of physical quantities in thermal equilibrium. This temperature dependence is contained in the functions W and U, and the symmetry properties of W and U give absolutely no clue to the temperature dependence of either. However, for our special purpose these symmetry properties are enough: we wish to prove a theorem true at all temperatures [Eq. (1.2)], and hence the proof need not involve the temperature explicitly.

7. MOMENT OF INERTIA IN QUANTUM STATISTICAL MECHANICS

Since the classical proof given in Sec. 4 is based upon a generalization of the equipartition theorem, and the equipartition theorem fails to hold in quantum statistical mechanics, our method of proof must be altered. We shall establish the following relation between the average value of L^2 and the mean square total linear momentum (in the x-direction, say):

$$\langle L^2\rangle_{\mathrm{Av}}=\langle(x_1^2+y_1^2)\rangle_{\mathrm{Av}}\langle P^2\rangle_{\mathrm{Av}}\times[1+\text{terms of order}(\Lambda S/V)].\tag{7.1}$$

The additional terms are not evaluated but tend to zero as the volume is made infinitely large in such a way that the ratio of surface area to volume becomes vanishingly small.[14] The proof depends upon the "correlation distance" Λ introduced in Sec. 5, and the surface terms are of relative order $\Lambda S/V$ where S is the surface area and V is the volume.

Once relation (7.1) is established, the result we want follows from (2.11), i.e., from the equipartition theorem for the center-of-gravity motion; this latter has been shown to hold in quantum statistical mechanics as well as in classical statistical mechanics.

Let us therefore proceed to prove (7.1). We write

$$l_i=p_{\theta i}/\hbar=-i(x_i\partial/\partial y_i-y_i\partial/\partial x_i),\tag{7.2}$$

and we use the permutation symmetry in order to write

$$\mathrm{Trace}(L^2e^{-\beta H})=N\hbar^2\,\mathrm{Trace}(l_1^2e^{-\beta H})+N(N-1)\hbar^2\,\mathrm{Trace}(l_1l_2e^{-\beta H}).\tag{7.3}$$

The traces are understood to be over all symmetric wave functions for Bose-Einstein statistics, over all antisymmetric wave functions for Fermi-Dirac statistics. We now use the results of the preceding section. We have from (6.4):

$$\begin{aligned}\mathrm{Trace}[l_1^2\exp(-\beta H)]&\\=\sum_{\mathbf{k}_1}\int\frac{d^3r_1}{V}&G(\mathbf{k}_1,\mathbf{r}_1)W(k_1^2)+\text{terms of order }\Lambda S/V\\=\sum_{\mathbf{k}_1}\int\frac{d^3r_1}{V}&(x_1^2k_{y1}^2+y_1^2k_{x1}^2-2x_1y_1k_{x1}k_{y1})W(k_1^2)\\&+\text{order }\Lambda S/V,\end{aligned}\tag{7.4}$$

---

[14] We should note that this condition excludes the possibility of making a cylindrical volume infinitely large by increasing the height of the cylinder without increasing its radius.

where the second line follows from the definition (6.3) of the function G. We introduce the notations

$$\begin{gathered}\langle x^2\rangle_{\mathrm{Av}}=V^{-1}\int x^2d^3r,\quad \langle y^2\rangle_{\mathrm{Av}}=V^{-1}\int y^2d^3r,\\ \langle xy\rangle_{\mathrm{Av}}=V^{-1}\int xyd^3r.\end{gathered}\tag{7.5}$$

It should be noted that these averages are *geometrical* averages, not averages over canonical ensembles. The geometrical averages arise because W is independent of position. With this notation we have

$$\begin{aligned}\mathrm{Trace}[l_1^2\exp(-\beta H)]=\sum_{\mathbf{k}_1}[&\langle x^2\rangle_{\mathrm{Av}}k_{y1}^2+\langle y^2\rangle_{\mathrm{Av}}k_{x1}^2\\ &-2\langle xy\rangle_{\mathrm{Av}}k_{x1}k_{y1}]W(k_1^2).\end{aligned}\tag{7.6}$$

We now use the fact that W depends on k^2 only:

$$\sum_{k_1}k_{x1}k_{y1}W(k_1^2)=0,\tag{7.7}$$

$$\begin{aligned}\sum_{k_1}k_{x1}^2W(k_1^2)&=\sum_{k_1}k_{y1}^2W(k_1^2)\\ &=\hbar^{-2}\,\mathrm{Trace}[p_{x1}^2\exp(-\beta H)],\end{aligned}\tag{7.8}$$

where the last part of Eq. (7.8) is true up to surface terms of order $\Lambda S/V$. We therefore get

$$\begin{aligned}N\hbar^2\,\mathrm{Trace}[l_1^2\exp(-\beta H)]&=N\langle(x^2+y^2)\rangle_{\mathrm{Av}}\\ &\times\mathrm{Trace}[p_{x1}^2\exp(-\beta H)]+\text{Order }\Lambda S/V.\end{aligned}\tag{7.9}$$

If this were classical mechanics, the trace of p_{x1}^2 would be determined by the equipartition theorem, and this would be the end of the proof. However, in quantum statistical mechanics the cross term [the last term of Eq. (7.3)] does not vanish, and must be taken into account explicitly. To do so, we use Eq. (6.22) and the properties of the function $U(\mathbf{k}_1,\mathbf{k}_2,\mathbf{r}_1,\mathbf{r}_2)$ which were established in the preceding section. According to the definition of $G(\mathbf{k}_1,\mathbf{k}_2,\mathbf{r}_1,\mathbf{r}_2)$, Eq. (6.19), we have for the operator $F_{12}=l_1l_2$:

$$\begin{aligned}G(\mathbf{k}_1,\mathbf{k}_2,\mathbf{r}_1,\mathbf{r}_2)=x_1x_2k_{y1}k_{y2}&+y_1y_2k_{x1}k_{x2}\\ &-x_1y_2k_{y1}k_{x2}-y_1x_2k_{x1}k_{y2}.\end{aligned}\tag{7.10}$$

Since (7.10) is already real, Eq. (6.22) shows that only the real part of U contributes.

Consider now the first term of (7.10); its contribution is

$$\sum_{\mathbf{k}_1,\mathbf{k}_2}\int\int\frac{d^3r_1d^3r_2}{V^2}x_1x_2k_{y1}k_{y2}U(\mathbf{k}_1,\mathbf{k}_2,\mathbf{r}_1,\mathbf{r}_2)\equiv A.\tag{7.11}$$

In order to simplify this expression, we first use the separability property (6.23). Consider values of $\mathbf{r}_1$ and $\mathbf{r}_2$ such that $|\mathbf{r}_1-\mathbf{r}_2|>\Lambda$; we then perform the sum over $\mathbf{k}$ vectors, and get zero:

$$\begin{aligned}\sum_{\mathbf{k}_1,\mathbf{k}_2}&k_{y1}k_{y2}U(\mathbf{k}_1,\mathbf{k}_2,\mathbf{r}_1,\mathbf{r}_2)\\ &\cong\Big(\sum_{\mathbf{k}_1}k_{y1}W(k_1^2)\Big)\Big(\sum_{\mathbf{k}_2}k_{y2}W(k_2^2)\Big)=0\\ &\qquad\text{for }|\mathbf{r}_1-\mathbf{r}_2|>\Lambda.\end{aligned}\tag{7.12}$$

We now introduce, in (7.11), the difference vector $\mathbf{r}_{12}=\mathbf{r}_2-\mathbf{r}_1$ as a variable of integration. Because of (7.12), large values of r_{12} make no contribution, hence we can extend the integration over the variable r_{12} over a volume V centered at $r_{12}=0$ (at $\mathbf{r}_1=\mathbf{r}_2$), rather than over the correct volume (which depends on $\mathbf{r}_1$); this procedure is correct except within a distance Λ of the walls, and hence the error made is of order $\Lambda S/V$. We thus write for the quantity A, (7.11):

$$\begin{aligned}A=\sum_{\mathbf{k}_1,\mathbf{k}_2}\int\frac{d^3r_1}{V}\int\frac{d^3r_{12}}{V}&\\ \times(x_1^2+x_1x_{12})&k_{y1}k_{y2}U(\mathbf{k}_1,\mathbf{k}_2,\mathbf{r}_{12}).\end{aligned}\tag{7.13}$$

We assert that the second term in the parenthesis gives zero upon integration. This follows from the inversion symmetry of U, Eq. (6.25); for consider the contribution made by $-\mathbf{k}_1$, $-\mathbf{k}_2$, $-\mathbf{r}_{12}$: the factor x_{12} changes sign, the factor $k_{y1}k_{y2}$ is unchanged, and so is U. Thus, in the integral over r_{12} and the sums over k_1 and k_2, we get pairwise cancellation of terms.

Thus we have

$$\begin{aligned}A&=\sum_{\mathbf{k}_1,\mathbf{k}_2}\int\int\frac{d^3r_1d^3r_2}{V^2}x_1x_2k_{y1}k_{y2}U(\mathbf{k}_1,\mathbf{k}_2,\mathbf{r}_1,\mathbf{r}_2)\\ &=\langle x^2\rangle_{\mathrm{Av}}\sum_{\mathbf{k}_1,\mathbf{k}_2}\int\frac{d^3r_{12}}{V}k_{y1}k_{y2}U(\mathbf{k}_1,\mathbf{k}_2,\mathbf{r}_{12})\\ &=\langle x^2\rangle_{\mathrm{Av}}\hbar^{-2}\,\mathrm{Trace}[p_{y1}p_{y2}\exp(-\beta H)]\\ &=\langle x^2\rangle_{\mathrm{Av}}\hbar^{-2}\,\mathrm{Trace}[p_{x1}p_{x2}\exp(-\beta H)],\end{aligned}\tag{7.14}$$

where the last step is a result of rotational invariance.

A similar reduction carried out on the second term of (7.10) gives the same as (7.14) but with $\langle x^2\rangle_{\mathrm{Av}}$ replaced by $\langle y^2\rangle_{\mathrm{Av}}$.

Finally, the last two terms of (7.10) make no contribution. Consider the third term of (7.10). We first replace x_1y_2 by x_1y_1 and use the inversion symmetry of U to show that the correction vanishes. Next we use translation symmetry in order to write:

$$\begin{aligned}\sum_{\mathbf{k}_1,\mathbf{k}_2}\int\int\frac{d^3r_1d^3r_2}{V^2}&x_1y_1k_{y1}k_{x2}U(\mathbf{k}_1,\mathbf{k}_2,\mathbf{r}_{12})\\ &=\langle xy\rangle_{\mathrm{Av}}\sum_{\mathbf{k}_1,\mathbf{k}_2}\int\frac{d^3r_{12}}{V}k_{y1}k_{x2}U(\mathbf{k}_1,\mathbf{k}_2,\mathbf{r}_{12}).\end{aligned}$$

To show that this is zero, we perform a simultaneous rotation of the vectors $\mathbf{k}_1$, $\mathbf{k}_2$, and $\mathbf{r}_{12}$ around the x-direction, through 180 degrees. The factor k_{y1} changes sign, the factor k_{x2} is unchanged, and U is also unchanged. Since the operation "sum over $\mathbf{k}_1$, $\mathbf{k}_2$ and integrate over $\mathbf{r}_{12}$" is invariant under such a rotation of dummy variables, we conclude that the result vanishes.

Thus we finally have, up to terms of relative order $\Lambda S/V$,

$$N(N-1)\hbar^2 \operatorname{Trace}[l_1 l_2 \exp(-\beta H)] = N(N-1) \times \langle (x^2+y^2)\rangle_{\text{Av}} \operatorname{Trace}[p_{x1}p_{x2}\exp(-\beta H)]. \quad (7.15)$$

Combining (7.3), (7.9), and (7.15), we obtain the desired result (7.1). This concludes our proof.

It should be noted that the proof given here applies to systems composed of identical particles. The extension to systems containing particles of several types is trivial and we shall not give it here. The crucial condition is the existence of a correlation length Λ which is independent of the volume of the container. The identity of the particles and the type of statistics obeyed by the particles is unessential; assuming particles of one kind only simplifies the mechanics of writing down the proof, and this is why we made this assumption.

8. CONCLUSION

We have now shown that the moment of inertia of a system in thermal equilibrium is equal to the usually computed quantity, both in classical statistical mechanics and in quantum statistical mechanics. There is, however, a considerable difference between the conditions of applicability of the two proofs. The classical proof is applicable to an arbitrary system in thermal equilibrium, with the only proviso being that the rotating motion must allow thermal equilibrium also, i.e., that the potential energy V of the system (excluding wall forces) is invariant under rotations about the axis which is being used as an axis of rotation.

The quantum mechanical proof, on the other hand, depended crucially upon the existence of a "correlation distance" Λ between particles with correlated momentum vectors. The reader may feel that this restriction is not essential, and is merely necessary for our particular proof. This is not so: there exists a system (the ideal Bose-Einstein gas below its condensation point) for which the moment of inertia I, defined by (3.9), is *not* equal to the usual value (1.2).[5] The ideal Bose-Einstein gas below the condensation point is characterized by a correlation distance equal to the linear dimensions of the container. This is so because a macroscopic number of particles are condensed into a state with wavelength equal to the dimensions of the container. Thus not only have we proved the theorem about the moment of inertia for finite Λ, but we can also produce a counter-example to show that no similar proof is possible if Λ can become infinite (i.e., as large as the container).

There are peculiar systems other than the ideal Bose-Einstein gas in which the correlation distance Λ is comparable to the linear dimensions of the container. For example, consider a pipe in the shape of a torus, filled with water, and assume that the water is flowing steadily in a clockwise direction. Then the momenta of particles in different parts of the pipe are correlated by the drift motion, and Λ is as large as the linear dimensions of the torus. Such a system is excluded for our purposes by the fact that it is not in thermal equilibrium if the walls of the pipe are at rest. We recall that the definition (3.9) of the moment of inertia involves averages over the thermal equilibrium distribution in a container at rest.

However, there are also some equilibrium systems with large correlation distances. For example, consider a container partially filled with water, in the complete absence of a gravitational field. This is a two-phase system, consisting of water and water vapor. Let us assume that the walls of the container are perfectly reflecting, so that the water does not adhere to the walls. In thermal equilibrium we then have one very large ball of water, with volume proportional to the volume of the container,[15] and this ball of water executes Brownian motion inside the container. This Brownian motion has a kinetic energy equal to $3kT/2$, and thus becomes slower and slower as the size of the system is increased. Nevertheless, there is then a correlation between momenta of particles anywhere within the ball of water, and the radius of the ball is proportional to the linear dimensions of the container. It is clear, also, that this system has very peculiar properties when it is set into rotation. Already at very low rotational velocities, the water ball breaks up and a sheath of water forms next to the walls, as a consequence of the centrifugal force. We have neglected the centrifugal force throughout (since we have always restricted ourselves to terms of order ω, and have ignored terms of order ω^2), and thus our whole analysis becomes unphysical for this case.

This is a fairly general property of many-phase systems in which the different phases can be separated in space by the centrifugal force. However we are primarily interested in applying our theorem to liquid helium, which is a single-phase system in which such things cannot happen.

These exceptional systems show that one must be careful to establish, in any particular case, that the conditions of validity of our theorem are satisfied. That is, one must show that a finite correlation distance Λ, independent of the volume of the container, does indeed exist. We shall do so for liquid helium, and we shall discuss the consequences of this theorem for the theory

---

[15] It is understood, of course, that the number of particles N is increased whenever the volume is increased, in such a way that the number density N/V stays constant.

of liquid helium, in the following paper. Our theorem also has some interesting applications for the theory of superconductivity, and these will be discussed separately.

ACKNOWLEDGMENTS

We would like to express our appreciation to Professor H. Messel who brought us together as a team, and whose constant encouragement and never failing initiative created the stimulating atmosphere to which this work is due.

APPENDIX: THE ROTATING CYLINDER

So far we have assumed that all linear dimensions of the container are large compared to the correlation length Λ. In applying these considerations to liquid helium, we shall need to consider a situation in which the "bucket" is a very flat cylinder, of radius $R \gg \Lambda$ but height D of the same order as Λ. It is intuitively obvious that the height of the cylinder is of no importance in thermal equilibrium, but it seems preferable to show formally that the proof given in Secs. 6 and 7 can be modified so as to allow for arbitrary values of D, and then leads to the result

$$I = I_0[1 + \text{order } (\Lambda/R)]. \tag{A.1}$$

The first modification is that we cannot use plane wave states of form $\exp[i(k_x x + k_y y + k_z z)]$, since the violation of the boundary conditions at $z=0$ and $z=D$ (the lower and upper plate of the cylinder, respectively) has serious consequences if $D \lesssim \Lambda$. However, we can use states of the form

$$\psi(\mathbf{k},\mathbf{r}) = \exp[i(k_x x + k_y y)] \sin(k_z z), \tag{A.2}$$

where $k_z = n\pi/D$, n integral. These functions obey the boundary conditions at $z=0$ and $z=D$, but violate the boundary condition at $r=R$. This latter violation is not serious if $R \gg \Lambda$, and leads to errors of the order of Λ/R.

The (unsymmetrized) product wave functions for the N-particle system are defined by

$$\Psi(\mathbf{K},\mathbf{R}) = \prod_{i=1}^{N} \psi(\mathbf{k}_i,\mathbf{r}_i), \tag{A.3}$$

where $\mathbf{K}$ is the ordered ensemble of the N vectors $\mathbf{k}_i$, and $\mathbf{R}$ the ordered ensemble of the N vectors $\mathbf{r}_i$, as before. In order to continue, we shall need notation for the projections of all these vectors on the x-y plane. We shall use Gothic letters for this purpose, e.g., the symbol $\mathfrak{r}_1$ is the projection on the x-y plane of the three-dimensional vector $\mathbf{r}_1$; i.e., $\mathfrak{r}_1$ has components $(x_1, y_1, 0)$. We then write (A.3) in the form

$$\Psi(\mathbf{K},\mathbf{R}) = \exp(i\mathfrak{K}\cdot\mathfrak{R}) \prod_{i=1}^{N} \sin(k_{zi} z_i). \tag{A.4}$$

We arrive at Eq. (6.11) just as before, but with $\exp(i\mathbf{K}\cdot\mathbf{R})$ replaced by $\Psi(\mathbf{K},\mathbf{R})$. At this stage, we restrict outselves to operators F_1 which do not act on the coordinate z_1 at all. We can therefore write

$$\begin{aligned} F_1\Psi(\mathbf{K},\mathbf{R}) &= \prod_{i=1}^{N} \sin(k_{zi} z_i) F_1 \exp(i\mathfrak{K}\cdot\mathfrak{R}) \\ &= \Psi(\mathbf{K},\mathbf{R}) \exp(-i\mathfrak{k}_1\cdot\mathfrak{r}_1) F_1 \exp(i\mathfrak{k}_1\cdot\mathfrak{r}_1) \\ &= \Psi(\mathbf{K},\mathbf{R}) G_1(\mathfrak{k}_1,\mathfrak{r}_1). \end{aligned} \tag{A.5}$$

The function G_1 introduced by (A.5) is the analog of the three-dimensional $G_1(\mathbf{k}_1,\mathbf{r}_1)$ defined by (6.12). With this definition of G_1, (6.13) is again correct (with the obvious replacement of $\exp(i\mathbf{K}\cdot\mathbf{R})$ by $\Psi(\mathbf{K},\mathbf{R})$ throughout). But in order to get a useful form, we now perform not only all the sums and integrations indicated in (6.14), but also a sum over k_{z1} and an integral over z_1. The two-dimensional analog of the function W, (6.14) is

$$\begin{aligned} W(\mathfrak{k}_1,\mathfrak{r}_1) = (N!)^{-1} \sum_{\mathrm{II}} \epsilon_{\mathrm{II}} \sum_{k_{z1}} \sum_{\mathbf{k}_2,\cdots,\mathbf{k}_N} \int \frac{dz_1}{D} \int \frac{d^3r_2 \cdots d^3r_N}{V^{N-1}} \\ \times \Psi^*(\mathbf{K},\mathbf{R}) \exp(-\beta H) \Psi(\mathbf{K},\mathbf{R}^{\mathrm{II}}) \end{aligned} \tag{A.6}$$

(6.15) is then replaced by

$$\begin{aligned} \mathrm{Trace}[F_1 \exp(-\beta H)] \\ = \tfrac{1}{2} \sum_{\mathfrak{k}_1} \int \frac{d^2r_1}{\pi R^2} [G_1^*(\mathfrak{k}_1,\mathfrak{r}_1) W(\mathfrak{k}_1,\mathfrak{r}_1) \\ + G_1(\mathfrak{k}_1,\mathfrak{r}_1) W^*(\mathfrak{k}_1,\mathfrak{r}_1)]. \end{aligned} \tag{A.7}$$

The subsequent proof that $W(\mathfrak{k}_1,\mathfrak{r}_1)$ is invariant under translations of the two-dimensional vector $\mathfrak{r}_1$, depends on $\mathfrak{k}_1$ only through $\mathfrak{k}_1^2 = k_{x1}^2 + k_{y1}^2$, and is real, goes through without difficulty. The function $G(\mathfrak{k}_1,\mathfrak{r}_1)$ is then defined to be the real part of $G_1(\mathfrak{k}_1,\mathfrak{r}_1)$, and Eq. (6.4) is replaced by

$$\begin{aligned} \mathrm{Trace}[F_1 \exp(-\beta H)] \\ = \sum_{\mathfrak{k}_1} \int \frac{d^2r_1}{\pi R^2} G(\mathfrak{k}_1,\mathfrak{r}_1) W(\mathfrak{k}_1^2)[1 + \mathrm{order}(\Lambda/R)]. \end{aligned} \tag{A.8}$$

The modification of the analysis of expectation values of two particle operators is completely similar. We replace vectors by their x-y plane projections systematically, and integrate over z_1, z_2 and sum over k_{z1}, k_{z2} before defining the two-dimensional analog of the function U, Eq. (6.20). The separability (6.23), translational invariance (6.24), and inversion invariance (6.25) have direct two-dimensional analogs. There is also rotation invariance under joint rotations of $\mathfrak{k}_1$, $\mathfrak{k}_2$, and $\mathfrak{r}_{12}$ about the z-direction. However, one invariance property is no longer possible: we cannot make a rotation by 180 degrees around the x-axis. Such a rotation leaves x-components invariant, and changes y- and z-components into their negatives. Since z-components are involved, there is no direct two-dimensional analog. However, we did not really need the change of sign of z-components for the proof. The two-dimensional symmetry operation which is needed is a mirroring in the x-z plane, which leaves x-components invariant, and changes y-components into their negatives. The two-dimensional U-function, $U(\mathfrak{r}_1,\mathfrak{r}_2,\mathbf{r}_{12})$, is invariant under this mirror operation.

From hereon the proof given in Sec. 7 goes through without change (except for the replacement of the 180 degree rotation about x by the mirroring operation). We thus obtain the desired result (A.1), in which the height of the cylinder does not enter into the estimate of error.

PHYSICAL REVIEW VOLUME 100, NUMBER 2 OCTOBER 15, 1955

Nonequilibrium Nature of the Superfluid State

S. T. Butler and J. M. Blatt
*The F. B. S. Falkiner Nuclear Research and Adolph Basser Computing Laboratories, School of Physics,**
The University of Sydney, Sydney, Australia

(Received April 4, 1955)

We show that the result of the preceding paper is directly applicable to liquid helium, i.e., no superfluid state can exist in liquid helium under conditions of thermodynamic equilibrium. The observed superfluid properties of liquid helium must therefore be due to a nonequilibrium (metastable) superfluid state. Previous theoretical arguments for equilibrium superfluidity are examined, and an experimental test of nonequilibrium superfluidity is suggested.

1. INTRODUCTION

Liquid helium below the λ point has superfluid properties, and many widely different theories have been advanced to account for this behavior. There is practically universal agreement in the theoretical literature, however, that the superfluid state of liquid helium is a state of thermal equilibrium.[1] The position is stated very clearly by Zilsel[2]: "The superflow is truly reversible; it is maintained not by the absence of collisions, but—under the given conditions—is the thermodynamic equilibrium state established by the collisions." Or, as Landau[3] puts it: "It is most essential that there is no friction between these two liquids moving through each other. . . . We get this relative motion when considering the statistical equilibrium in a uniformly rotating vessel. But if there can be some sort of relative motion in the state of statistical equilibrium, it means that it cannot be accompanied by friction." London[4] remarks that this point is still open to question as far as experimental confirmation is concerned, but considers the evidence for truly reversible (i.e., equilibrium) superflow extremely strong.

We have shown[5] that the ideal Bose-Einstein gas below its transition point does indeed have a peculiar relationship between the thermal equilibrium value $\bar{L}(\omega)$ of the angular momentum and the angular velocity ω of the container. In particular, the moment of inertia I, defined by

$$I=\lim_{\omega\to 0}(\partial\bar{L}/\partial\omega) \qquad (1.1)$$

is less than the classical value, in spite of the fact that the system is in thermal equilibrium. However, a more detailed discussion showed[5] that the system would appear to be "infrafluid" in actual experiments, because the experiments average over rather wide ranges of the angular velocity ω. It was proposed that the weakly interacting Bose-Einstein gas might serve as a model of a *nonequilibrium* superfluid, that is, under certain reasonable assumptions about collision cross sections, the superfluid state would be metastable with a lifetime longer than the times used in the relevant experiments. This idea has been elaborated since by Klemens.[6]

The investigation of the behavior of the ideal Bose-Einstein gas suggested strongly that the superfluid state of liquid helium is not an equilibrium state, that it is maintained by the absence (or rather, the scarcity) of collisions, rather than being the thermodynamic equilibrium state established by the collisions. How-

---

* Also supported by the Nuclear Research Foundation within the University of Sydney.

[1] The nearest we could find to the contrary is in a paper by N. F. Mott, Phil. Mag. **40**, 61 (1949). Mott first reviews arguments for an energy gap between the ground state and all excited states of an interacting Bose-Einstein gas, and then goes on to discuss superfluidity in terms of a metastable state. However, if there really were an energy gap, the superfluid state would be thermodynamically stable. Thus Mott's discussion of the rotating toroid is inconsistent with his earlier arguments for an energy gap.

[2] P. R. Zilsel, Phys. Rev. **92**, 1106 (1953).

[3] L. Landau, J. Phys. (U.S.S.R.) **5**, 71 (1941).

[4] F. London, *Superfluids* (John Wiley and Sons, Inc., New York, 1954), Vol. 2.

[5] J. M. Blatt and S. T. Butler, this issue [Phys. Rev. **100**, 476 (1955)]. See also Phys. Rev. **96**, 1149 (1954).

[6] P. G. Klemens (to be published).