Elementary statistical mechanics of a relativistic gas in thermal equilibrium

1974 J. Phys. A: Math. Nucl. Gen. 7 2130

(http://iopscience.iop.org/0301-0015/7/17/008)

J. Phys. A: Math., Nucl. Gen., Vol. 7, No. 17, 1974. Printed in Great Britain.

Elementary statistical mechanics of a relativistic gas in thermal equilibrium

K Kraus† and P T Landsberg‡

† Physikalisches Institut der Universität Würzburg, Würzburg, Germany

‡ Department of Mathematics, University of Southampton, Southampton, England

Received 22 March 1974, in final form 12 May 1974

Abstract. Starting from an assumption of ergodicity, the macroscopic description of a relativistic gas in terms of an energy–momentum tensor and four-currents for different particle types is obtained by summation over particles and time averaging. As a result of relativistic particle kinematics the formalism becomes covariant, and ergodicity as defined here is shown to be a Lorentz invariant concept.

1. Introduction

Special relativistic statistical mechanics is a discipline which has developed only recently. Nevertheless so many papers have been devoted to it already that we can cite only a selection (Møller 1968, van Kampen 1969, Nakajima 1969, Balescu and Brenig 1971, Stewart 1971). We want to present here some elementary considerations concerning the simplest possible system: a gas of particles confined to a box. We found it advantageous to follow the ergodic approach, thus defining time-independent (ie, equilibrium) quantities which describe the gas macroscopically by means of time averaging. The methods thus developed should be applicable to more interesting situations, too.

The main point of the paper is that by using the new idea of a state indicator which is attached to the box enclosing the system, and which can be read by any observer, the 'state of a system' becomes a rather clear Lorentz invariant concept (§ 2). In the usual way of looking at this problem the state of a system is a more complicated idea. This arises from the relativity of simultaneity which leads to different observers combining the motions of the constituent particles differently to arrive at their idea of 'the state of the system'. The preferred frame of reference provided by the rest frame of the box is thus given due importance in our work. However, one must still sum over the particles, taking relativistic kinematics correctly into account. This is achieved by the kinematic weight factors of § 3. The emergence of familiar results in this section (eg equations (3.4) and (3.5)) by matching macroscopic and microscopic methods thus leads to confidence in our approach. The definition of the rest frame I_0, and of ergodicity, are discussed in the concluding two sections.

The discussion of a particular case of our model (with fixed particle number) and the comparison of our method with previous work of one of us (Landsberg 1970, Landsberg and Johns 1970a, b) will be the subject of a separate investigation.

2. Time-based probabilities and averages in the rest frame

Consider a gas contained in a box which is at rest in an inertial frame I_0 (in a sense to be specified more precisely in § 4) and occupies in I_0 the volume V^0. A state i of the gas is, for our purpose, sufficiently characterized by specifying the momenta $\boldsymbol{p}_{ir}^0$ and energies ϵ_{ir}^0 $(r = 1, \ldots, N_i)$ of the gas particles r, the total number N_i of which may depend on the state i. The superfix 0 indicates that we are using the rest frame I_0. With a suitably coarse-grained momentum space, $\boldsymbol{p}_{ir}^0$ and ϵ_{ir}^0 may be restricted to have discrete values. Since only a finite total energy is available to the gas particles, the total number K of states i (distinct with respect to the coarse graining) which may occur during the time development of the gas is also finite.

The internal state i of the gas may be determined in principle by an observer in I_0, at every time t^0, simply by measuring the momenta and energies of all particles in the box at that time. Due to the finiteness of signal velocities, the totality of this information will actually reach the observer at some later time $t^0 + \Delta t^0$, with a time lag Δt^0 of order $1/c$ times the diameter l^0 of the box. However, Δt^0 can be arranged to be independent of t^0 and the gas state i, and drops out if one considers only time differences (which is all we need to do). Imagine, for simplicity, that all data about gas particles are fed into a computer, which then at every time t^0 indicates the state i of the gas at time $t^0 - \Delta t^0$. This device, called state indicator in the following, shall be also at rest in I_0. We will assume that the transition times between different states i are very short as compared to the life times of the states themselves, such that at almost all times t^0 the state indicator reports a definite state i. The state changes are due to collisions of the gas particles among themselves and with the walls of the box. Such events need not be separated by space-like intervals, so that the life time of a state i may well be considerably shorter than l^0/c.

The main (ergodicity) assumption of this approach is the following (see also § 5). Assume the state indicator to be observed in I_0 during a time interval τ^0, and consider the total life time τ_i^0 of a particular state i which is the sum of all time intervals $\Delta\tau_i^0$ for which the reading of the state indicator was i. Then

$$\lim_{\tau^0 \to \infty} \frac{\tau_i^0}{\tau^0} \equiv q_i^0 \tag{2.1}$$

is assumed to exist, and to be independent of the state from which the system is started off initially, for all states i. The quantity q_i^0 is the time-based probability for the occurrence of the state i in the rest frame I_0.

For inertial frames I different from I_0 the gas is also defined to be in the state shown on the state indicator. This leads to a considerable simplification in our treatment compared with most other attempts: state changes are simply point events on the straight world line of the state indicator. The states i discussed here thus do not refer to momenta and energies of single particles at fixed times t in an inertial frame, unless that frame be I_0. If one considers that the information which leads to an entry on the state indicator is gathered from a sampling surface t^0 = constant in I_0, then this is also the sampling surface for all other frames I in which use is made of the state indicator to determine the state of the system.

A first conclusion follows immediately. If the state indicator is observed from another inertial frame I with respect to which I_0 has a velocity $\boldsymbol{w}$, the total time interval of observation and the total life time of the state i are $\tau = \gamma\tau^0$ and $\tau_i = \gamma\tau_i^0$, respectively, with $\gamma = (1 - w^2/c^2)^{-1/2}$. Defining in the frame I time-based probabilities for the state i, in analogy to (2.1), by

$$\lim_{\tau \to \infty} \frac{\tau_i}{\tau} = q_i, \tag{2.2}$$

one finds

$$q_i = q_i^0. \tag{2.3}$$

Since therefore the probabilities q_i^0 are Lorentz invariant, the index 0 can be omitted.

Quantities of physical interest can be obtained for a fixed state i from quantities belonging to single particles r with values Q_{ir}^0 in I_0 by summation over all particles. Examples are the total momentum

$$\boldsymbol{P}_i^0 = \sum_{r=1}^{N_i} \boldsymbol{p}_{ir}^0$$

and the total energy

$$E_i^0 = \sum_{r=1}^{N_i} \epsilon_{ir}^0$$

of the gas in state i in I_0. The averages of such quantities over long times are taken to characterize the gas macroscopically. In the frame I_0 these quantities are obtained in an obvious way. For example,

$$\langle \boldsymbol{P}^0 \rangle_0 = \sum_{i=1}^{K} q_i \sum_{r=1}^{N_i} \boldsymbol{p}_{ir}^0 \tag{2.4}$$

and

$$\langle E^0 \rangle_0 = \sum_{i=1}^{K} q_i \sum_{r=1}^{N_i} \epsilon_{ir}^0 \tag{2.5}$$

are the average momentum and energy of the gas in I_0. The index 0 at the brackets indicates that the time averaging is performed in I_0.

The frame I_0 is specified here by the condition

$$\langle \boldsymbol{P}^0 \rangle_0 = 0. \tag{2.6}$$

Some physical motivation for this condition will be given in § 4. The expression (3.4) derived in § 3 for the average momentum in an arbitrary inertial frame implies that I_0 is indeed the only frame in which the gas has zero average momentum.

Besides (2.6) we make an *assumption of spatial homogeneity*, which may be formulated as follows. Consider a single particle r belonging to state i of the gas. For a sufficiently long time interval τ^0 in I_0, this particle exists during the total time $\tau_i^0 = q_i \tau^0$, according to the definition of q_i. We now assume, in addition, that the particle spends the fraction $\Delta V^0/V^0$ of its total life time τ_i^0 in any given part ΔV^0 of the total gas volume V^0 in I_0. In other words, the time-based proability for the localization of any particle is uniformly distributed throughout V^0.

Under this assumption it makes sense to define, as time averages, densities for various physical quantities. The energy density is

$$\rho^0 = \frac{1}{V^0} \sum_{ir} q_i \epsilon_{ir}^0 = \frac{1}{V^0} \langle E^0 \rangle_0, \tag{2.7}$$

whereas the momentum density is $(1/V^0)\langle \boldsymbol{P}^0 \rangle_0$ and vanishes according to (2.6). Moreover, transport properties may be discussed in the usual way. For instance, with the particle velocities

$$\boldsymbol{u}^0_{ir} = \frac{c^2 \boldsymbol{p}^0_{ir}}{\epsilon^0_{ir}}, \tag{2.8}$$

the average momentum transported by the particles through a surface element dS^0 with unit normal $\boldsymbol{e}^0$ during a time interval dt^0 becomes

$$\frac{1}{V^0} \sum_{ir} q_i(\boldsymbol{e}^0 \cdot \boldsymbol{u}^0_{ir}) \boldsymbol{p}^0_{ir} \, dS^0 \, dt^0. \tag{2.9}$$

The argument leading to (2.9) is a standard one in statistical mechanics. If particle r belonging to state i crosses dS^0 in the time interval between t^0 and $t^0 + dt^0$, it transports the momentum $\boldsymbol{p}^0_{ir}$ through dS^0. This will happen if, and only if, the particle is localized at time t^0 somewhere inside a cylinder with base dS^0 and height $\boldsymbol{e}^0 \cdot \boldsymbol{u}^0_{ir} \, dt^0$, the volume of which is $\Delta V^0 = |\boldsymbol{e}^0 \cdot \boldsymbol{u}^0_{ir}| \, dS^0 \, dt^0$ (see figure 1). If $\boldsymbol{e}^0 \cdot \boldsymbol{u}^0_{ir} < 0$, the particle crosses dS^0 in a direction opposite to $\boldsymbol{e}_0$, in which case $\boldsymbol{p}^0_{ir}$ has to be counted with a minus sign.

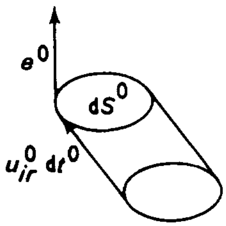

Figure 1.

According to the above homogeneity assumption, the probability of finding the particle inside ΔV^0 at the arbitrarily chosen instant t^0 is $q^i(\Delta V^0/V^0)$. Thus the particle gives the contribution

$$\frac{1}{V^0} q_i(\boldsymbol{e}^0 \cdot \boldsymbol{u}^0_{ir}) \boldsymbol{p}^0_{ir} \, dS^0 \, dt^0$$

to the average momentum transport through dS^0. Summation over particles r and states i leads to (2.9). We should mention, however, that in this derivation we have excluded from consideration the possibility that some state change affects the particle r in the above cylinder just between the times t^0 and $t^0 + dt^0$. This is justified here since we have neglected anyway the time required for state changes as compared to the life times of the states i.

We now add an *isotropy assumption* by requiring that the average momentum transport (2.9) in the rest frame I_0 gives rise to an isotropic pressure p. Since for a gas with pressure p the quantity corresponding to (2.9) is $p\boldsymbol{e}^0 \, dS^0 \, dt^0$, this yields the equation

$$p\boldsymbol{e}^0 = \frac{1}{V^0} \sum_{ir} q_i(\boldsymbol{e}^0 \cdot \boldsymbol{u}^0_{ir}) \boldsymbol{p}^0_{ir}, \tag{2.10}$$

valid for all unit vectors $\boldsymbol{e}^0$. Equation (2.10) also implies

$$p = \frac{1}{V^0} \sum_{ir} q_i(\boldsymbol{e}^0 \cdot \boldsymbol{u}_{ir}^0)(\boldsymbol{e}^0 \cdot \boldsymbol{p}_{ir}^0) \tag{2.11}$$

$$= \frac{1}{3V^0} \sum_{ir} q_i(\boldsymbol{u}_{ir}^0 \cdot \boldsymbol{p}_{ir}^0), \tag{2.12}$$

where (2.12) follows from (2.11) by summation over three mutually orthogonal unit vectors $\boldsymbol{e}^0$.

In the same way, the energy flux through $\mathrm{d}S^0$ with unit normal $\boldsymbol{e}^0$ in a time interval $\mathrm{d}t^0$ if averaged over time yields

$$\frac{1}{V^0} \sum_{ir} q_i(\boldsymbol{e}^0 \cdot \boldsymbol{u}_{ir}^0)\epsilon_{ir}^0 \, \mathrm{d}S^0 \, \mathrm{d}t^0 = \frac{c^2}{V^0} \sum_{ir} q_i(\boldsymbol{e}^0 \cdot \boldsymbol{p}_{ir}^0) \, \mathrm{d}S^0 \, \mathrm{d}t^0 = 0$$

according to (2.8) and (2.6). Thus in I_0 the average energy current is zero.

Collecting the quantities obtained so far, one may define an energy–momentum tensor $T_{\mu\nu}^0$ of the gas in the rest frame I_0 as

$$T_{\mu\nu}^0 = \mathrm{diag}(p, p, p, \rho^0) \tag{2.13}$$

inside the box volume V^0 and zero outside.

Since particles of different kind may be present in the gas, we will finally introduce densities and currents for particles of a given type. The various kinds of particles will be distinguished by an index s, and we introduce for any state i the symbol

$$\delta_{rs}^{(i)} = \begin{cases} 1 & \text{if the particle } r \text{ in state } i \text{ is of type } s, \\ 0 & \text{otherwise.} \end{cases}$$

The total number of particles of type s in state i is then $\Sigma_{r=1}^{N_i} \delta_{rs}^{(i)}$, and by averaging over time one finds the mean particle number of type s in the rest frame I_0 as

$$\langle N_s \rangle_0 = \sum_{ir} q_i \delta_{rs}^{(i)}. \tag{2.14}$$

The assumption of spatial homogeneity implies that it also makes sense to introduce the corresponding density

$$\rho_s^0 = \frac{1}{V^0} \sum_{ir} q_i \delta_{rs}^{(i)} = \frac{1}{V^0} \langle N_s \rangle_0. \tag{2.15}$$

By the same assumption, the average number of particles of type s crossing a surface element $\mathrm{d}S^0$ with unit normal $\boldsymbol{e}^0$ in a time interval $\mathrm{d}t^0$ is

$$\langle \mathrm{d}N_s \rangle_0 = \frac{1}{V^0} \sum_{ir} q_i \delta_{rs}^{(i)}(\boldsymbol{e}^0 \cdot \boldsymbol{u}_{ir}^0) \, \mathrm{d}S^0 \, \mathrm{d}t^0,$$

and we assume this to be zero for each $\boldsymbol{e}^0$ and s, which is equivalent to

$$\sum_{ir} q_i \delta_{rs}^{(i)} \boldsymbol{u}_{ir}^0 = 0. \tag{2.16}$$

The macroscopic situation with respect to the density and flow of any particle type s may then be described in I_0 by a four-current density ($\nu = 1, \ldots, 4$)

$$j_{s\nu}^0 = \{\boldsymbol{0}, c\rho_s^0\} \tag{2.17}$$

inside and zero outside V^0. The four-divergence of $j^0_{s\nu}$ is identically zero. This is not true for the energy–momentum tensor $T^0_{\mu\nu}$ as defined by (2.13), due to non-vanishing contributions from the boundary of V^0 which must be present since the gas alone, without the walls, would not be a stable system.

3. Kinematic weight factors and time averages for a general inertial frame

Consider an interial frame I with respect to which the box has velocity $\boldsymbol{w}$. We want to define time averages for the gas as seen from I in the same way, as in § 2 for the frame I_0. The quantities Q^0_{ir} occurring there, which refer to a particle r in state i, have to be replaced by Q_{ir}, which are the quantities Q^0_{ir} when transformed from I_0 to I.

Since by (2.3) the time-based probabilities q_i for the occurrence of a state i are the same in all frames I, one might be tempted to define time averages in I by the expression

$$\sum_{i=1}^{K}\sum_{r=1}^{N_i} q_i Q_{ir}. \tag{3.1}$$

This would lead to an average momentum and an average energy

$$\sum_{ir} q_i \boldsymbol{p}_{ir} \qquad \text{and} \qquad \sum_{ir} q_i \epsilon_{ir}, \tag{3.2}$$

respectively, of the gas in I.

However, (3.1) and (3.2) are incorrect, as easily seen. Since the particles r composing the gas in state i move freely, $\{\boldsymbol{p}_{ir}, (1/c)\epsilon_{ir}\}$ transform as four-vectors, while the q_i are scalars. Hence (3.2) implies that the average momentum and energy of the gas form also a four-vector. This, however, is not acceptable (Landsberg and Johns 1967). One way to show this is as follows. In the rest frame I_0 the average energy–momentum tensor $T^0_{\mu\nu}$ of the gas is given by (2.13). In the frame I one obtains from it by Lorentz transformation the tensor

$$T_{\mu\nu} = (\rho^0 + p)\frac{u_\mu u_\nu}{c^2} + p\eta_{\mu\nu} \tag{3.3}$$

with the four-velocity

$$u_\nu = \gamma\{\boldsymbol{w}, c\}, \qquad \gamma = \left(1 - \frac{w^2}{c^2}\right)^{-1/2}$$

of the box in I and the Minkowski metric

$$\eta_{\mu\nu} = \text{diag}(1, 1, 1, -1).$$

More precisely, $T_{\mu\nu}$ is given by (3.3) inside the world tube occupied by the moving box and zero outside. Since the volume of the box in I is $V = \gamma^{-1}V^0$, one obtains from (3.3) with the momentum density $(1/c)\{T_{14}, T_{24}, T_{34}\}$ and energy density T_{44} the formulae

$$\begin{aligned}\langle \boldsymbol{P}\rangle &= V\frac{\gamma^2}{c^2}(\rho^0 + p)\boldsymbol{w} = \frac{\gamma}{c^2}\boldsymbol{w}(\rho^0 V^0 + pV^0)\\ &= \frac{\gamma}{c^2}\boldsymbol{w}(\langle E^0\rangle_0 + pV^0)\end{aligned} \tag{3.4}$$

and

$$\langle E\rangle = V[\gamma^2(\rho^0+p)-p] = \gamma V^0\rho^0+\frac{V^0}{\gamma}(\gamma^2-1)p$$

$$= \gamma\left(\langle E^0\rangle_0+\frac{w^2}{c^2}pV^0\right) \tag{3.5}$$

for the average momentum and energy in I, respectively. Equations (3.4) and (3.5) may be re-expressed by stating that, instead of $\{\langle \boldsymbol{P}\rangle, (1/c)\langle E\rangle\}$, the quantity

$$\left\{\langle \boldsymbol{P}\rangle, \frac{1}{c}(\langle E\rangle+pV)\right\} \tag{3.6}$$

is a four-vector (Landsberg and Johns 1967). In I_0 the corresponding quantity is

$$\left\{\langle \boldsymbol{P}^0\rangle_0, \frac{1}{c}(\langle E^0\rangle_0+pV^0)\right\} = \left\{\boldsymbol{0}, \frac{1}{c}(\langle E^0\rangle_0+pV^0)\right\}. \tag{3.7}$$

The four-vector transformation law applied to (3.7) leads back to (3.4) and (3.5). The four-vector property of (3.6) is connected with the fact that $T_{\mu\nu}$ is not divergence-free throughout space–time, for otherwise $\{\langle \boldsymbol{P}\rangle, (1/c)\langle E\rangle\}$ would have to be a four-vector.

A simple argument will now be given in order to show that (3.1) is indeed not the correct formula for time averages in I. Assume the gas, as seen from I_0, to be in state i during some time interval $\Delta\tau_i^0$. Then, by definition of the state i, the particle r with momentum $\boldsymbol{p}_{ir}^0$ and energy ϵ_{ir}^0 exists during that time interval in I_0. Its velocity is

$$\boldsymbol{u}_{ir}^0 = \frac{c^2\boldsymbol{p}_{ir}^0}{\epsilon_{ir}^0},$$

so that it will travel during $\Delta\tau_i^0$ the distance

$$\Delta\boldsymbol{x}_{ir}^0 = \boldsymbol{u}_{ir}^0\Delta\tau_i^0 \tag{3.8}$$

in I_0. In I, it thus exists during the time interval

$$\Delta\tau_{ir} = \gamma\left(\Delta\tau_i^0+\frac{\boldsymbol{w}\cdot\Delta\boldsymbol{x}_{ir}^0}{c^2}\right) = \gamma\Delta\tau_i^0\left(1+\frac{\boldsymbol{w}\cdot\boldsymbol{u}_{ir}^0}{c^2}\right) \tag{3.9}$$

resulting from $\{\Delta\boldsymbol{x}_{ir}^0, c\Delta\tau_i^0\}$ with (3.8) by a Lorentz transformation. Thus, although in I the system as a whole, according to the reading of the state indicator, is found in the state i during the time interval

$$\Delta\tau_i = \gamma\Delta\tau_i^0, \tag{3.10}$$

the corresponding life times (3.9) of the individual particles r constituting this state i differ from (3.10) by the *kinematic weight factors*

$$1+\frac{\boldsymbol{w}\cdot\boldsymbol{u}_{ir}^0}{c^2}. \tag{3.11}$$

A correct evaluation of time averages in I has to contain these weight factors, and therefore leads to the expression

$$\langle Q\rangle = \sum_{i=1}^{K} q_i \sum_{r=1}^{N_i}\left(1+\frac{\boldsymbol{w}\cdot\boldsymbol{u}_{ir}^0}{c^2}\right)Q_{ir} \tag{3.12}$$

instead of (3.1). This is found to lead to a consistent macroscopic description of the gas in terms of time averages in an arbitrary inertial frame I.

One can first check that the correct transformation properties (3.4) and (3.5) for the average momentum and energy result from (3.12). The only assumption needed for this is equation (2.6). If seen from I, the momentum, energy and velocity of the particle r are, respectively,

$$\boldsymbol{p}_{ir} = \boldsymbol{p}_{ir}^0 + \boldsymbol{w}\left((\gamma-1)\frac{\boldsymbol{w}\cdot\boldsymbol{p}_{ir}^0}{w^2} + \frac{\gamma\epsilon_{ir}^0}{c^2}\right), \tag{3.13}$$

$$\epsilon_{ir} = \gamma(\epsilon_{ir}^0 + \boldsymbol{w}\cdot\boldsymbol{p}_{ir}^0), \tag{3.14}$$

$$\begin{aligned}\boldsymbol{u}_{ir} &= \frac{c^2\boldsymbol{p}_{ir}}{\epsilon_{ir}} \\ &= \left(1+\frac{\boldsymbol{w}\cdot\boldsymbol{u}_{ir}^0}{c^2}\right)^{-1}\left[\gamma^{-1}\boldsymbol{u}_{ir}^0 + \boldsymbol{w}\left(1+(1-\gamma^{-1})\frac{\boldsymbol{w}\cdot\boldsymbol{u}_{ir}^0}{w^2}\right)\right].\end{aligned} \tag{3.15}$$

Applying (3.12) with $Q_{ir} = \boldsymbol{p}_{ir}$ as given by (3.13), we get

$$\begin{aligned}\langle \boldsymbol{P}\rangle &= \sum_{ir} q_i\left(1+\frac{\boldsymbol{w}\cdot\boldsymbol{u}_{ir}^0}{c^2}\right)\left[\boldsymbol{p}_{ir}^0 + \boldsymbol{w}\left((\gamma-1)\frac{\boldsymbol{w}\cdot\boldsymbol{p}_{ir}^0}{w^2}+\frac{\gamma\epsilon_{ir}^0}{c^2}\right)\right] \\ &= \sum_{ir} q_i\left[\boldsymbol{p}_{ir}^0 + \boldsymbol{w}(\gamma-1)\frac{\boldsymbol{w}\cdot\boldsymbol{p}_{ir}^0}{w^2} + \frac{\gamma}{c^2}\boldsymbol{w}\epsilon_{ir}^0 + \frac{\boldsymbol{w}\cdot\boldsymbol{u}_{ir}^0}{c^2}\boldsymbol{p}_{ir}^0\right. \\ &\quad \left. + \frac{\gamma-1}{c^2}\boldsymbol{w}\left(\frac{\boldsymbol{w}}{w}\cdot\boldsymbol{u}_{ir}^0\right)\left(\frac{\boldsymbol{w}}{w}\cdot\boldsymbol{p}_{ir}^0\right) + \frac{\gamma}{c^2}\boldsymbol{w}\left(\boldsymbol{w}\cdot\frac{\epsilon_{ir}^0\boldsymbol{u}_{ir}^0}{c^2}\right)\right].\end{aligned}$$

By (2.6) and (2.8), the first, second and last term are zero. The third term gives $(\gamma/c^2)\boldsymbol{w}\langle E^0\rangle_0$. By (2.10) and (2.11) with $\boldsymbol{e}^0 = \boldsymbol{w}/w$, the fourth and fifth term give $(pV^0/c^2)\boldsymbol{w}$ and $[(\gamma-1)pV^0/c^2]\boldsymbol{w}$, respectively. Thus, finally,

$$\langle \boldsymbol{P}\rangle = \frac{\gamma}{c^2}\boldsymbol{w}(\langle E^0\rangle_0 + pV^0),$$

in accordance with (3.4). Similarly we get from (3.12) and (3.14)

$$\begin{aligned}\langle E\rangle &= \sum_{ir} q_i\left(1+\frac{\boldsymbol{w}\cdot\boldsymbol{u}_{ir}^0}{c^2}\right)\gamma(\epsilon_{ir}^0 + \boldsymbol{w}\cdot\boldsymbol{p}_{ir}^0) \\ &= \gamma\sum_{ir} q_i\left[\epsilon_{ir}^0 + \boldsymbol{w}\cdot\boldsymbol{p}_{ir}^0 + \boldsymbol{w}\cdot\frac{\epsilon_{ir}^0\boldsymbol{u}_{ir}^0}{c^2} + \frac{w^2}{c^2}\left(\frac{\boldsymbol{w}}{w}\cdot\boldsymbol{u}_{ir}^0\right)\left(\frac{\boldsymbol{w}}{w}\cdot\boldsymbol{p}_{ir}^0\right)\right].\end{aligned}$$

As above, the second and third term vanish, whereas the first term gives $\gamma\langle E^0\rangle_0$ and the last one $\gamma(w^2/c^2)pV^0$. The result coincides with (3.5).

The transformation properties (3.4) and (3.5) have been obtained above by applying a Lorentz transformation to the energy–momentum tensor $T_{\mu\nu}^0$. The latter was defined in § 2 to describe in I_0 the energy–momentum densities and corresponding transport properties of the gas, averaged with respect to time. An analogous interpretation must be possible in I for the transformed energy–momentum tensor $T_{\mu\nu}$. One easily convinces oneself that the assumption of spatial homogeneity as formulated in § 2 allows the definition of average densities and transport properties also in the inertial frame I. Formulae like (2.7) for the energy density or (2.9) for the momentum transport remain valid in I if properly modified, which simply means dropping the index 0 everywhere, and inserting for all particles r the kinematic weight factors (3.11). We have thus to compare the components of $T_{\mu\nu}$ with such time averages in order to establish consistency of the formalism.

Equation (3.3) and the familiar physical interpretation of the components of $T_{\mu\nu}$ yield the following macroscopic quantities as describing the gas in I:

(i) An energy density

$$T_{44} = \gamma^2\rho^0+(\gamma^2-1)p = \gamma^2\left(\rho^0+\frac{w^2}{c^2}p\right). \tag{3.16}$$

(ii) A momentum density

$$\frac{1}{c}\{T_{14}, T_{24}, T_{34}\} = \frac{\gamma^2}{c^2}\boldsymbol{w}(\rho^0+p). \tag{3.17}$$

(iii) An energy flux density $c\{T_{41}, T_{42}, T_{43}\}$, such that the energy flowing in a time interval $\mathrm{d}t$ through a surface element $\mathrm{d}S$ with unit normal $\boldsymbol{e}$ is

$$\gamma^2(\rho^0+p)(\boldsymbol{w}\,.\,\boldsymbol{e})\,\mathrm{d}S\,\mathrm{d}t. \tag{3.18}$$

(iv) Flux densities $\{T_{k1}, T_{k2}, T_{k3}\}$ for the kth component of momentum ($k = 1, 2, 3$), such that the momentum flowing in $\mathrm{d}t$ through $\mathrm{d}S$ with normal $\boldsymbol{e}$ is

$$\left(p\boldsymbol{e}+\frac{\gamma^2}{c^2}\boldsymbol{w}(\rho^0+p)(\boldsymbol{w}\,.\,\boldsymbol{e})\right)\,\mathrm{d}S\,\mathrm{d}t. \tag{3.19}$$

From the microscopic point of view, the quantities corresponding to (3.16)–(3.19) are:

(i) $$\frac{1}{V}\sum_{ir} q_i\left(1+\frac{\boldsymbol{w}\,.\,\boldsymbol{u}_{ir}^0}{c^2}\right)\epsilon_{ir} = \frac{1}{V}\langle E\rangle,$$

(ii) $$\frac{1}{V}\sum_{ir} q_i\left(1+\frac{\boldsymbol{w}\,.\,\boldsymbol{u}_{ir}^0}{c^2}\right)\boldsymbol{p}_{ir} = \frac{1}{V}\langle \boldsymbol{P}\rangle,$$

(iii) $$\frac{1}{V}\sum_{ir} q_i\left(1+\frac{\boldsymbol{w}\,.\,\boldsymbol{u}_{ir}^0}{c^2}\right)(\boldsymbol{e}\,.\,\boldsymbol{u}_{ir})\epsilon_{ir}\,\mathrm{d}S\,\mathrm{d}t,$$

(iv) $$\frac{1}{V}\sum_{ir} q_i\left(1+\frac{\boldsymbol{w}\,.\,\boldsymbol{u}_{ir}^0}{c^2}\right)(\boldsymbol{e}\,.\,\boldsymbol{u}_{ir})\boldsymbol{p}_{ir}\,\mathrm{d}S\,\mathrm{d}t.$$

Comparing (i) and (ii) with (3.16) and (3.17), respectively, we obtain equations (3.4) and (3.5) (with a factor $V^{-1} = \gamma(V^0)^{-1}$), which have been checked already. With $\epsilon_{ir}\boldsymbol{u}_{ir}/c^2 = \boldsymbol{p}_{ir}$, the microscopic quantity (iii) becomes

$$\begin{aligned}
&\frac{c^2}{V}\sum_{ir} q_i\left(1+\frac{\boldsymbol{w}\,.\,\boldsymbol{u}_{ir}^0}{c^2}\right)(\boldsymbol{e}\,.\,\boldsymbol{p}_{ir})\,\mathrm{d}S\,\mathrm{d}t\\
&\qquad = \frac{c^2}{V}(\boldsymbol{e}\,.\,\langle\boldsymbol{P}\rangle)\,\mathrm{d}S\,\mathrm{d}t\\
&\qquad = \frac{\gamma^2}{V^0}(\boldsymbol{e}\,.\,\boldsymbol{w})(\langle E^0\rangle_0+pV^0)\,\mathrm{d}S\,\mathrm{d}t\\
&\qquad = \gamma^2(\rho^0+p)(\boldsymbol{w}\,.\,\boldsymbol{e})\,\mathrm{d}S\,\mathrm{d}t,
\end{aligned}$$

in accordance with (3.18). Likewise, after inserting $\boldsymbol{p}_{ir}$ and $\boldsymbol{u}_{ir}$ from (3.13) and (3.15) in the microscopic quantity (iv), an elementary calculation reproduces (3.19).

As a last application of (3.12), consider the four-current density of particles of a given type s in the inertial frame I. By Lorentz transformation of j^0_{sv} given by (2.17), we find in I the four-current density

$$j_{sv} = \{\gamma \rho^0_s \boldsymbol{w}, \gamma c \rho^0_s\} \tag{3.20}$$

inside and zero outside the moving volume V. From this macroscopic point of view, the average number of s particles in the gas is

$$\langle N_s \rangle = V \frac{1}{c} j_{s4} = \frac{V^0}{\gamma} \gamma \rho^0_s \equiv \langle N_s \rangle_0, \tag{3.21}$$

and the average number of particles of type s crossing a surface element $\mathrm{d}S$ with unit normal $\boldsymbol{e}$ in a time interval $\mathrm{d}t$ is

$$\langle \mathrm{d}N_s \rangle = (\boldsymbol{e} \cdot \gamma \rho^0_s \boldsymbol{w})\, \mathrm{d}S\, \mathrm{d}t. \tag{3.22}$$

The equality of $\langle N_s \rangle$ and $\langle N_s \rangle_0$ comes from the fact that j_{sv} is a conserved current.

Again this macroscopic description is consistent with the microscopic point of view, ie the microscopic definitions

$$\langle N_s \rangle = \sum_{ir} q_i \left(1 + \frac{\boldsymbol{w} \cdot \boldsymbol{u}^0_{ir}}{c^2}\right) \delta^{(i)}_{rs}$$

and

$$\langle \mathrm{d}N_s \rangle = \frac{1}{V} \sum_{ir} q_i \left(1 + \frac{\boldsymbol{w} \cdot \boldsymbol{u}^0_{ir}}{c^2}\right) (\boldsymbol{e} \cdot \boldsymbol{u}_{ir}) \delta^{(i)}_{rs}\, \mathrm{d}S\, \mathrm{d}t$$

immediately reproduce, with (2.16) and (3.15), the expressions (3.21) and (3.22). The decisive condition is now (2.16), which corresponds to the role of condition (2.6) for energy and momentum.

4. Discussion of the rest frame condition (2.6)

Assume that box and gas together form a closed system, and denote by $\boldsymbol{R}^0_i$ and F^0_i momentum and energy, respectively, of the box alone in I_0 when the gas state is i. Then the total momentum

$$\mathscr{P}^0 = \boldsymbol{P}^0_i + \boldsymbol{R}^0_i \tag{4.1}$$

and the total energy

$$\mathscr{E}^0 = E^0_i + F^0_i \tag{4.2}$$

are independent of i, and the same is true for the centre of mass velocity

$$\boldsymbol{w}^0 = \frac{c^2}{\mathscr{E}^0} \mathscr{P}^0 \tag{4.3}$$

of the total system. As the total system has to be strictly at rest in I_0, we have $\boldsymbol{w}^0 = 0$ and thus $\mathscr{P}^0 = 0$. After time averaging, (4.1) thus yields

$$\langle \boldsymbol{P}^0 \rangle_0 + \langle \boldsymbol{R}^0 \rangle_0 = 0 \tag{4.4}$$

with the average momentum of the box

$$\langle \boldsymbol{R}^0 \rangle_0 = \sum_i q_i \boldsymbol{R}_i^0. \tag{4.5}$$

Equation (4.4) characterizes the rest frame I_0 uniquely. For our formalism, however, we need a characterization which contains quantities referring to the gas (without the box) only. We want to show now that equation (2.6) is appropriate for this purpose.

It is indeed very plausible (eg, from spatial isotropy or reflection invariance arguments) that in most cases equation (4.4) will be satisfied only if both $\langle \boldsymbol{P}^0 \rangle_0$ and $\langle \boldsymbol{R}^0 \rangle_0$ vanish separately. In order to discuss this in some more detail, consider the centre of mass of the gas alone, which in state i in I_0 has the coordinates

$$\boldsymbol{X}_i^0(t^0) = \frac{1}{E_i^0} \sum_r \epsilon_{ir}^0 \boldsymbol{x}_{ir}^0(t^0) \tag{4.6}$$

(with the coordinates $\boldsymbol{x}_{ir}^0(t^0)$ of the freely moving particles r), and which is moving in state i with the constant velocity

$$\dot{\boldsymbol{X}}_i^0 = \frac{c^2}{E_i^0} \boldsymbol{P}_i^0. \tag{4.7}$$

According to our general assumptions, any state change $i \to j$ in I_0 occurs instantaneously, at time $t^{0\prime}$ say. Assume that such state change is due to either

(i) a collision of two or more particles r belonging to state i at some collision point $\boldsymbol{x}^{0\prime}$ inside the box, producing some other particles s belonging to state j at the same point $\boldsymbol{x}^{0\prime}$, or

(ii) a collision with a wall of the box of one or more particles r belong to state i, which are either absorbed, or produce some other particles s belonging to state j at the point of collision.

It is easily seen from (4.6), (4.7) and the conservation or energy and momentum in the collision process that neither $\boldsymbol{X}^0$ nor $\dot{\boldsymbol{X}}^0$ are changed during state transitions of type (i). A state change (ii), however, in general leads to discontinuous changes $\boldsymbol{X}_i^0(t^{0\prime}) \to \boldsymbol{X}_j^0(t^{0\prime})$ and $\dot{\boldsymbol{X}}_i^0 \to \dot{\boldsymbol{X}}_j^0$ of both $\boldsymbol{X}^0$ and $\dot{\boldsymbol{X}}^0$.

If one assumes, however, that collisions of gas particles with the walls are always elastic, then (4.6) together with energy conservation for the colliding particles immediately implies that $\boldsymbol{X}^0$ does not jump even in type (ii) transitions: $\boldsymbol{X}_j^0(t^{0\prime}) = \boldsymbol{X}_i^0(t^{0\prime})$. In this particular case, the centre of mass of the gas thus describes a continuous path $\boldsymbol{X}^0(t^0)$ in I_0. The total distance travelled by it during a time interval τ^0 is

$$\tau^0 \sum_i (q_i + \Delta_i) \dot{\boldsymbol{X}}_i^0 \tag{4.8}$$

with some correction terms Δ_i which vanish for $\tau^0 \to \infty$, according to the definition of the probabilities q_i. This distance obviously cannot exceed in length the diameter of the box. Dividing (4.8) by τ_0 and taking the limit $\tau^0 \to \infty$ thus yields

$$\sum_i q_i \dot{\boldsymbol{X}}_i^0 = 0 \tag{4.9}$$

or, with (4.7) and $E_i^0 \equiv E^0$ (independent of i) since no energy is exchanged with the walls,

$$\frac{c^2}{E^0} \sum_i q_i \boldsymbol{P}_i^0 = \frac{c^2}{E^0} \langle \boldsymbol{P}^0 \rangle_0 = 0,$$

ie, condition (2.6).

But even if inelastic collisions of gas particles with the walls are permitted, equation (2.6) is still to be expected to hold in I_0, at least as a very good approximation, for a large class of systems. Assume, for the following, that the energy changes ΔE^0 due to inelastic collisions are very small as compared to the total energy E_i^0 of the gas in any state i. This is true for almost all cases of physical interest, since E_i^0 contains the rest energies of all gas particles. Now consider a collision of gas particles with the wall, occurring at some point $\boldsymbol{x}^0$, during which an energy ΔE^0 is transferred to the box. If E^0 and $\boldsymbol{X}^0$ denote the total energy and the centre of mass coordinates of the gas before the collision, (4.6) yields for the centre of mass coordinates after the collision:

$$\boldsymbol{X}^0 + \Delta\boldsymbol{X}^0 = \frac{1}{E^0 - \Delta E^0}(E^0\boldsymbol{X}^0 - \Delta E^0\boldsymbol{x}^0).$$

With $\Delta E^0/E^0 \equiv \alpha \ll 1$ we get, neglecting quadratic and higher order terms in α,

$$\Delta\boldsymbol{X}^0 = \alpha(\boldsymbol{X}^0 - \boldsymbol{x}^0).$$

Thus any discontinuous jump of the centre of mass in I_0 is at most of the order of magnitude αl^0, with l^0 the box diameter in I_0 and $\alpha \ll 1$. Moreover, such jumps are not expected to occur in some preferred direction, but instead to nearly compensate each other during a sufficiently long period of observation, and thus to be negligible for the overall centre of mass motion of the gas. This leads again to equation (4.9).

By (4.7) and (4.9),

$$\sum_i q_i \frac{c^2}{E_i^0}\boldsymbol{P}_i^0 = 0.$$

Putting

$$E_i^0 = \langle E^0\rangle_0 + \Delta E_i^0 = \langle E^0\rangle_0(1+\beta_i)$$

with $\beta_i = \Delta E_i^0/\langle E^0\rangle_0 \ll 1$ (according to our previous assumption) and

$$\sum_i q_i\beta_i = 0, \tag{4.10}$$

this leads to

$$0 = \frac{c^2}{\langle E^0\rangle_0}\sum_i q_i\boldsymbol{P}_i^0(1+\beta_i)^{-1} = \frac{c^2}{\langle E^0\rangle_0}\left(\sum_i q_i\boldsymbol{P}_i^0 - \sum_i q_i\beta_i\boldsymbol{P}_i^0\right),$$

if quadratic and higher order terms in β_i are again neglected. If one also drops the linear terms, (2.6) is again recovered. This drastic procedure, however, is not necessary, since the average $\Sigma_i\, q_i\beta_i\boldsymbol{P}_i^0$ is expected to be nearly zero anyway. Otherwise the momenta $\boldsymbol{P}_i^0$ and their weights $q_i\beta_i$ (constrained by (4.10)) in this average would have to be particularly correlated, which is very implausible for the irregular type of motion of the gas particles considered here.

Due to the collisions with the gas particles, the box vibrates and, as a whole, performs an irregular 'Brownian' motion in I_0. For a heavy box with sufficiently massive walls, however, neither its vibrations nor the motion of its centre of mass will be macroscopically visible. It is thus justified, as we have done here, to consider the box as strictly at rest and of constant shape and volume in I_0.

5. Ergodicity

Following the usual scheme of ergodic theory (Farquhar 1964), we will now try to re-obtain the time averages discussed so far in an alternative way, namely, as averages over a suitable ensemble of systems which are macroscopically identical copies of the single system considered before. Let us first consider such an ensemble in the rest frame I_0.

We imagine a very large number $\mathcal{N}$ of such systems, all at rest in I_0, which are selected at random in the sense that the selection procedure does not depend on the microscopic states i of these systems at the time of selection $t^{0\prime}$. Then, if each one of these systems evolves in time in the way discussed in § 2 and $\mathcal{N}$ is large enough, just $\mathcal{N}q_i$ of them are expected to be in state i at time $t^{0\prime}$. Moreover, by the same argument, this distribution of states i in the ensemble will be preserved in time, ie, at any other time $t^0 \neq t^{0\prime}$ one will also find $\mathcal{N}q_i$ systems (but in general different ones) in the state i. In this way the *time-based probabilities* q_i also determine, quite naturally, the *relative frequencies* of the states i in the ensemble at every fixed time of observation t^0.

But even more follows from the above assumption of random selection. First, the distribution of states i in the ensemble is expected to be given by the probabilities q_i not only for fixed times t^0 in I_0, but on any other space-like hypersurface as well, eg, a space-like hyperplane corresponding to a fixed time t in some other inertial frame I. If we imagine, as in § 2, that to each system of the ensemble a state indicator is attached, then this means that the relative frequencies for the different readings i of these state indicators in the ensemble are given by q_i also at any fixed time t in an arbitrary inertial frame I. Thus the probabilities q_i are *Lorentz invariant* also with respect to their role as relative frequencies of states i in an ensemble of identical systems.

Second, assume one is looking at the ensemble, in an arbitrary frame I, more closely by taking into account at time t not only the state indicator readings i but also the occurrence of the individual particles r which constitute the states i. Again we expect, from the complete randomness of the microscopic situation, that at any time t the relative frequency of each particle r in the total ensemble will be found to agree with the time-based probability for the occurrence of the same particle r in the history of a single system in I. The latter, according to § 3, is q_i times the kinematic weight factor (3.11).

Thus if we finally define, for quantities to which the particle r in state i gives the contribution Q_{ir} as discussed before, *ensemble averages* at any fixed time t in an arbitrary inertial frame I by summing over all particles r in the single systems which are present at that time t and averaging over the total ensemble, we obtain

$$\sum_{ir} q_i \left(1 + \frac{\boldsymbol{w} \cdot \boldsymbol{u}_{ir}^0}{c^2}\right) Q_{ir},$$

which coincides with the formula (3.12) for *time averages* of a single system. In particular, in the rest frame I_0 we recover the time averages $\Sigma_{ir}\, q_i Q_{ir}^0$ used in § 2.

This argument makes it plausible that the time averages for a single system are indeed representative for the average properties of a suitable ensemble of systems. Thus our systems share with ergodic ones the property that time averages and ensemble averages coincide. Therefore we called our crucial hypotheses (2.1) an ergodicity assumption. Moreover, because time averages and ensemble averages coincide in all inertial frames I. our systems can be considered to be *ergodic in a Lorentz invariant sense.*

The correspondence between our ergodicity condition and the usual ones should however not be taken too literally. We will illustrate this for the particular case in which the total energy E^0 of the gas in I_0 is conserved. Clearly, then, all states $i = 1, \ldots, K$ occurring in our formalism belong to the same energy. But no further specification of these states or the probabilities q_i is required here, whereas in ergodic theory the set of possible states is usually taken to cover the whole energy shell, and the q_i are derived from Lebesgue measure in phase space. Accordingly, our formalism includes a description of systems which are *not* ergodic in the most narrow sense of the word.

If, for instance, besides E^0 there exists another constant of motion C, then the states $i = 1, \ldots, K$ occurring in our formalism belong to fixed E^0 *and* C, thus forming a proper subset of all energetically possible states. Now assume C to be a macroscopic quantity, such that states i with different values C_i are macroscopically distinguishable. Then a selection of 'macroscopically identical' systems will produce an ensemble of systems with a fixed value of C as well. If now assumption (2.1) is true for sets of states $i = 1, \ldots, K$ with given E^0 and C, our previous reasoning applies without any modification.

But even if C is not macroscopic, the situation is not entirely hopeless. In this case, a macroscopically selected ensemble necessarily contains systems with different C values. Since, however, C is now a 'microscopic' quantity, it might happen that ensemble averages for 'macroscopic' quantities like pressure, density etc are not changed too much if one replaces the macroscopically selected ensemble (with different C values) by any one of its sub-ensembles belonging to fixed C. Such sub-ensembles, however, again fit into our formalism.

The final aim of our theory is a 'macroscopic' description of a given *single* system. The ensemble averages considered so far may be used for this purpose only if, in the representing ensemble, the mean square deviations of all quantities considered are sufficiently small. As in ordinary statistical mechanics, this will be true for large systems (ie, with large particle numbers) only. Ensemble averages then correspond, as well known, to the properties of such systems in thermal equilibrium.

References

Balescu R and Brenig L 1971 *Physica* **54** 504–21

Farquhar I E 1964 *Ergodic Theory in Statistical Mechanics* (London, New York, Sydney: Interscience)

van Kampen N G 1969 *Physica* **43** 244–62

Landsberg P T 1970 *Essays in Physics* vol 2 (London: Academic Press) pp 93–129

Landsberg P T and Johns K A 1967 *Nuovo Cim.* B **52** 28–44

—— 1970a *J. Phys.* A: *Gen. Phys.* **3** 113–20

—— 1970b *Proc. Int. Conf. on Thermodynamics, Cardiff* 1970 (London: Butterworths) pp 487–92

Møller C 1968 *K. Danske Vid. Selsk. Mat.-Fys. Medd.* **36** No 16

Nakajima S 1969 *Prog. Theor. Phys.* **41** 1450–60

Stewart J M 1971 *Non-Equilibrium Relativistic Kinetic Theory* (Lecture Notes in Physics vol 10) (Berlin, Heidelberg, New York: Springer)