Elementary statistical mechanics of a relativistic gas in thermal equilibrium. II

1977 J. Phys. A: Math. Gen. 10 953

(http://iopscience.iop.org/0305-4470/10/6/015)

J. Phys. A: Math. Gen., Vol. 10, No. 6, 1977. Printed in Great Britain. © 1977

Elementary statistical mechanics of a relativistic gas in thermal equilibrium II

P T Landsberg† and K A Johns‡

† Department of Mathematics, University of Southampton, Southampton, SO9 5NH, UK
‡ Department of Medical Statistics, Welsh National School of Medicine, Cardiff, UK

Received 13 April 1976, in final form 8 December 1976

Abstract. Elementary methods of relativistic statistical mechanics developed in a previous paper are applied to the particular case of a relativistic gas with constant particle number. They are compared with an earlier approach and are shown to be inequivalent and superior to it. A third approach, based on a radical re-appraisal of the principles underlying the earlier one, is introduced. The two new schemes are examined, and the difficulties inherent in each are considered.

1. Introduction

In a previous paper a formulation of the special relativistic statistical mechanics of a large but finite, ergodic system of particles in a box was given (Kraus and Landsberg 1974, to be referred to as A). Particle creation and annihilation were allowed, and the particles were of course able to collide. Macroscopic quantities were obtained in a general inertial frame I in terms of the momenta $\boldsymbol{p}_{ir}$ and energies ϵ_{ir} of individual particles $r = 1, 2 \ldots N_i$ in a general state i of the system by summation over particles and time averaging over states, using the 'time-based' probabilities q_i of finding the system in a state i. Such macroscopic quantities were, e.g., the average energy $\langle E \rangle$ and the average momentum $\langle \boldsymbol{P} \rangle$ of the gas and its pressure p. The Lorentz transformation properties of the macroscopic averages were then shown to follow correctly from relativistic particle kinematics.

The simplifying feature of this work was the assumption that a 'state indicator', attached to the wall of the box, enables any observer to read off immediately the state of the system. In this way we avoided complicated Lorentz transformation properties of the 'state of the system' which would result if an observer moving relative to the box would characterize a state by the momenta and energies of individual particles as measured simultaneously in his own rest frame. Instead only the effect of the observer's motion on the time intervals during which he sees the different state indicator readings had to be taken into account. Just as one can argue that in thermodynamics the rest frame of a system plays a preferred rôle (Landsberg 1970), so one can argue that this also holds in statistical mechanics. The most reasonable way whereby a moving observer can determine the state of a system is to have appropriate measuring devices and dials attached to the system, and then to utilize their readings.

For other recent discussions of relativistic thermodynamics see, e.g., Habeger 1972, Grøn 1973, Szamosi 1973. We have not established a connection as yet between the present work and relativistic kinetic theory (e.g., Anderson and Witting 1974, and references cited therein) although such a connection should exist.

In this paper we shall first discuss a particular case of model treated in A, namely, a gas with constant particle number ($N_i = N$, independent of i). We shall then compare the approach of A with an older attempt to formulate an elementary version of relativistic statistical mechanics (Landsberg and Johns 1970a, to be referred to as B). The notation of A will be retained. Note in particular that a subscript or superscript 0, attached to a symbol for a quantity measured in a general inertial frame I, means that it is measured in the rest frame I_0 of the box. The probability that the system is in state i will be denoted by q_i, a Lorentz invariant quantity ($q_i = q_i^0$) by equation (2.3) of A. The total lifetime of a state i during a period of observation τ is denoted by τ_i so that (cf equation (2.2) of A)

$$q_i = \lim_{\tau\to\infty} \frac{\tau_i}{\tau}, \qquad \sum_i q_i = 1.$$

The velocity of the box in I is denoted by $\boldsymbol{w}$. Strictly speaking, $\boldsymbol{w}$ should be interpreted as the centre of mass velocity of the total system (box plus particles), since it is the latter which is constant in I whereas the box velocity undergoes some small fluctuations (cf. § 4 of A). Since, however, the centre of mass remains inside the box for all times, $\boldsymbol{w}$ is also the *average* box velocity in all inertial frames I.

Finally, we shall introduce a new approach to statistical mechanics by considering not the states in which a system exists for limited periods of time, but rather the events at which it changes from one state to another. We shall show equivalence between this new procedure and the more usual one, and demonstrate its particular suitability for use under the Lorentz transformation. It will thus be apparent that both the theory of A, based on particle states, and an augmented version of the theory of B, based on states of the whole system, are valid in appropriate conditions. A comparison is made between both theories, and the conceptual difficulties brought about by each are discussed.

2. The model without particle creation or destruction

In this special case a given particle will be denoted by the same label r in all states i of the system. The distance $\boldsymbol{d}_r$ travelled by particle r during a time interval $\tau = \Sigma_i \tau_i$ (with τ_i equal to the lifetime of state i) in a frame I may differ from the distance $\boldsymbol{w}\tau$ travelled by the box of some vector $\boldsymbol{l}_r$ whose length, however, cannot exceed the diameter of the box in I:

$$\boldsymbol{w}\tau + \boldsymbol{l}_r = \boldsymbol{d}_r = \sum_i \tau_i \left(1 + \frac{\boldsymbol{w} \cdot \boldsymbol{u}_{ir}^0}{c^2}\right) \boldsymbol{u}_{ir}.$$

In the explicit expressions for $\boldsymbol{d}_r$ on the right-hand side, the 'kinematical weight factors' (equation (3.11) of A) have been taken into account. Dividing by τ and passing to the limit $\tau \to \infty$ yields (cf (2.2) of A):

$$\langle \boldsymbol{u}_r \rangle \equiv \sum_i q_i \left(1 + \frac{\boldsymbol{w} \cdot \boldsymbol{u}_{ir}^0}{c^2}\right) \boldsymbol{u}_{ir} = \boldsymbol{w} \qquad (r = 1, 2 \ldots N). \tag{2.1}$$

In particular, the time averaged velocity of any particle vanishes in I_0:

$$\langle \boldsymbol{u}_r^0 \rangle_0 = \sum_i q_i \boldsymbol{u}_{ir}^0 = 0 \qquad (r = 1, 2 \ldots N). \tag{2.2}$$

(Note that (2.2) and (3.15) of A together yield an alternative derivation of (2.1).)

It is an immediate consequence of (2.2) that if one rewrites equation (3.12) of A as

$$\langle Q \rangle = \sum_{ir} \pi_{ir} Q_{ir}, \qquad \pi_{ir} = q_i \left(1 + \frac{\boldsymbol{w} \cdot \boldsymbol{u}_{ir}^0}{c^2}\right), \tag{2.3}$$

then the π_{ir} are normalized in the sense

$$\sum_i \pi_{ir} = 1 \qquad (r = 1, 2 \ldots N), \qquad \sum_{ir} \pi_{ir} = N. \tag{2.4}$$

The π_{ir} are the time-based probabilities that a given particle r contributes to state i, thus having energy-momentum $\{\boldsymbol{p}_{ir}, \epsilon_{ir}/c\}$, in frame I. (The total probability that particle r has given values $\{\boldsymbol{p}, \epsilon/c\}$ of energy-momentum in I is thus the sum of the π_{ir} over all states i with $\boldsymbol{p}_{ir} = \boldsymbol{p}$ and $\epsilon_{ir} = \epsilon$.) Thus it is formally possible to use averaging procedures which utilize the π_{ir}. In passing from the rest frame I_0 to a general frame I they transform according to

$$\frac{\pi_{ir}}{\pi_{ir}^0} = 1 + \frac{\boldsymbol{w} \cdot \boldsymbol{u}_{ir}^0}{c^2}, \qquad \pi_{ir}^0 \equiv q_i. \tag{2.5}$$

If one introduces an average particle velocity in state i,

$$\boldsymbol{u}_i = \frac{1}{N} \sum_r \boldsymbol{u}_{ir}, \tag{2.6}$$

and a particle-averaged probability

$$\pi_i = \frac{1}{N} \sum_r \pi_{ir}, \tag{2.7}$$

then by summing over r in (2.5) one finds an equation from which the particle label r has been eliminated altogether:

$$\frac{\pi_i}{\pi_1^0} = 1 + \frac{\boldsymbol{w} \cdot \boldsymbol{u}_i^0}{c^2}, \qquad \pi_1^0 = q_i. \tag{2.8}$$

These new probabilities are also normalized,

$$\sum_i \pi_i = 1, \tag{2.9}$$

as follows from (2.4) and (2.7). Also $\Sigma_i q_i \boldsymbol{u}_i^0 = 0$ by (2.2) and (2.6), and this leads again to (2.9) from (2.8). In our theory the π_{ir} can be used for averaging in all frames, but the π_i can be used for averaging in frame I_0 only where there are just the q_i (equations (2.4) and (2.5) of A).

3. Comparison with previous work

The preceding section enables us to make a comparison with the previous paper B mentioned in the introduction. We shall deal first with the similarities between A and B.

In B it was assumed that the state i of the system could be determined from any inertial frame I (but it was not explained how this might be done), and probabilities π_i for the system to be in states $i = 1, 2 \ldots$ were defined in terms of the ratio of two terms as in equation (2.2) of A. This immediately led to a Lorentz transformation formula of the form (2.8) for the π_i; but whereas (2.8) contains the average particle velocity $\boldsymbol{u}_i^0$ in I_0, the corresponding velocity in B was interpreted as the centre of mass velocity of the gas in I_0. As we shall see later on, the method used in B to discuss states of the system without taking into account in detail the individual particles constituting these states does not reproduce the results of the approach described in A, and therefore must be rejected. Nonetheless both approaches may be derived in an analogous way from a common source, as follows.

Suppose on treats the π_i as unknown probabilities that in frame I the gas is in state i, characterized by a total momentum $\boldsymbol{P}_i$ and a total energy E_i of the gas. Then we require

$$\sum_i \pi_i = 1, \qquad \sum_i \pi_i E_i = \langle E \rangle, \qquad \sum_i \pi_i \boldsymbol{P}_i = \langle \boldsymbol{P} \rangle \tag{3.1}$$

and try to find the quantities f_i in the assumed form

$$\pi_i = (1 + f_i)\pi_i^0 \tag{3.2}$$

of the transformation law which connects the π_i in I with the corresponding probabilities π_i^0 in the rest frame I_0 (in which $\langle \boldsymbol{P}^0 \rangle_0 = 0$). The quantities $\{\boldsymbol{P}_i, E_i/c\}$ on the left-hand sides of (3.1) are four-vectors (they refer to periods in which all gas particles move freely, as explained in A and B). The quantities $\langle \boldsymbol{P} \rangle$ and $\langle E \rangle$ on the right-hand sides, however, form the four-vector $\{\langle \boldsymbol{P} \rangle, (\langle E \rangle + pV)/c\}$ which contains, in addition, the box volume V in I and the pressure p. Therefore the π_i in (3.1) must change in some way under Lorentz transformations, i.e., $f_i \neq 0$ in (3.2). Relating π_i to π_i^0 by (3.2) and $\{\boldsymbol{P}_i, E_i/c\}$ to $\{\boldsymbol{P}_i^0, E_i^0/c\}$ via Lorentz transformation, the circumstance that the transformation laws for $\langle \boldsymbol{P} \rangle$ and $\langle E \rangle$ are known yields the following equations for f_i (Landsberg and Johns 1970b):

$$\sum_i \pi_i^0 f_i E_i^0 = 0, \tag{3.3}$$

$$\sum_i \pi_i^0 f_i \boldsymbol{P}_i^0 = \frac{\boldsymbol{w}}{c^2} p V^0. \tag{3.4}$$

Since $\langle P^0 \rangle_0 = 0$, (3.3) is satisfied by

$$f_i = \frac{\boldsymbol{w} \cdot \boldsymbol{P}_i^0}{E_i^0} \equiv \frac{\boldsymbol{w} \cdot \boldsymbol{u}_i^0}{c^2}. \tag{3.5}$$

Then (3.4) and (3.5) yield the relation used in B for the pressure p:

$$pV^0 \boldsymbol{n} = \sum_i \pi_i^0 (\boldsymbol{n} \cdot \boldsymbol{u}_i^0) \boldsymbol{P}_i^0, \tag{3.6}$$

where $\boldsymbol{n}$ is an arbitrary unit vector. With (3.5) we have obtained a transformation of the π_i of the form (2.8) in which $\boldsymbol{u}_i^0$, however, means the centre of mass velocity $c^2 P_i^0/E_i^0$ of the gas in I_0. This essentially leads to the results of the earlier attempt B.

What was not known then is that a completely analogous approach can be set up which involves sums over particles. Consider π_{ir} as unknown probabilities that a given particle r appears as a constituent of state i. thus having energy-momentum $\{\boldsymbol{p}_{ir}, \epsilon_{ir}/c\}$, in frame I. Instead of (3.1) and (3.2) we then require

$$\sum_i \pi_{ir} = N, \qquad \sum_{ir} \pi_{ir}\epsilon_{ir} = \langle E\rangle, \qquad \sum_{ir} \pi_{ir}\boldsymbol{p}_{ir} = \langle \boldsymbol{P}\rangle \tag{3.7}$$

and

$$\pi_{ir} = (1+f_{ir})\pi_i^0, \tag{3.8}$$

where $\pi_{ir}^0 = \pi_i^0$ (independent of r) since, by definition, the gas is in a given state in I_0 if and only if each particle r has an energy-momentum $\{\boldsymbol{p}_{ir}^0, \epsilon_{ir}^0/c\}$ characteristic for that state. (The π_i^0 have to be identified with the $q_i^0 = q_i$ used in A.) The two equations for the unknown f_{ir} corresponding to (3.3) and (3.4) now turn out to be

$$\sum_{ir} \pi_i^0 f_{ir}\epsilon_{ir}^0 = 0, \tag{3.9}$$

$$\sum_{ir} \pi_i^0 f_{ir}\boldsymbol{p}_{ir}^0 = \frac{\boldsymbol{w}}{c^2} pV^0. \tag{3.10}$$

Equation (3.9) is satisfied by

$$f_{ir} = \frac{\boldsymbol{w} \cdot \boldsymbol{p}_{ir}^0}{\epsilon_{ir}^0} = \frac{\boldsymbol{w} \cdot \boldsymbol{u}_{ir}^0}{c^2} \tag{3.11}$$

which leads at once to our theory A, with a pressure p given (from (3.10) and (3.11)) by

$$pV^0\boldsymbol{n} = \sum_{ir} \pi_i^0(\boldsymbol{n} \cdot \boldsymbol{u}_{ir}^0)\boldsymbol{p}_{ir}^0. \tag{3.12}$$

Thus the constraint that both $\{\boldsymbol{P}_i, E_i/c\}$ and $\{\langle\boldsymbol{P}\rangle, (\langle E\rangle + pV)/c\}$ shall be four-vectors is satisfied by the theories of A *and* B, and is thus not sufficient to single out one of them.

We shall finally prove that the theories of A and B are in fact inequivalent. It suffices to show that equations (3.6) and (3.12) yield different values of the pressure p for one and the same system. We use the formulae:

$$3V^0c^2p = \sum_i \pi_i^0 \frac{(\boldsymbol{P}_i^0)^2}{E_i^0} \tag{3.13}$$

and

$$3V^0c^2p = \sum_i \pi_i^0 \sum_r \frac{(\boldsymbol{P}_{ir}^0)^2}{\epsilon_{ir}^0} \tag{3.14}$$

which are equivalent to (3.6) and (3.12), respectively (cf, for example, (2.11) and (2.12) of A). Their inequivalence is easily shown as follows. Let $\boldsymbol{a}_1 \ldots \boldsymbol{a}_N$ be arbitrary vectors and $c_1 \ldots c_N$ arbitrary positive numbers with $\Sigma_r c_r = 1$. Then

$$\left(\sum_r c_r\boldsymbol{a}_r\right)^2 \leqslant \sum_r c_r(\boldsymbol{a}_r)^2, \tag{3.15}$$

and equality holds in (3.15) if and only if $\boldsymbol{a}_1 = \boldsymbol{a}_2 = \ldots = \boldsymbol{a}_N$. This follows immediately from Cauchy's inequality

$$\left(\sum_r \alpha_r\beta_r\right)^2 \leqslant \sum_r \alpha_r^2 \sum_r \beta_r^2$$

(for which equality holds if and only if $\beta_r = a\alpha_r$, $r = 1 \ldots N$) by taking $\alpha_r = \sqrt{c_r}$, $\beta_r = a_{rk}\sqrt{c_r}$ (with $a_{rk} = k$-th component of $\boldsymbol{a}_r$) and adding the resulting inequalities for $k = 1, 2$, and 3. Applying (3.15) with $\boldsymbol{a}_r = \boldsymbol{u}_{ir}^0/c^2 = \boldsymbol{p}_{ir}^0/\epsilon_{ir}^0$ and $c_r = \epsilon_{ir}^0/E_i^0$ leads to

$$\frac{(\boldsymbol{P}_i^0)^2}{E_i^0} \leqslant \sum_r \frac{(\boldsymbol{p}_{ir}^0)^2}{\epsilon_{ir}^0}$$

for all i. Thus (3.13) always yields a pressure p which is smaller than or at most equal to the pressure calculated from (3.14). Moreover, the pressures are equal if and only if the particle velocities $\boldsymbol{u}_{ir}^0$ are independent of r for all states i, a case which is without any practical interest. The attempt in B to obtain the momentum transport across a surface as a fraction of the total momentum $\boldsymbol{P}_i^0$, transported through the surface by the centre of mass motion of the gas, thus underestimates the pressure. This is, after all, not unexpected, since the momenta $\boldsymbol{p}_{ir}^0$ of individual particles, which all give a positive contribution in (3.14), compensate each other to a large extent in the total momentum $\boldsymbol{P}_i^0$ which enters (3.13).

From the foregoing and the observation that in both theories $\{\langle \boldsymbol{P}\rangle, (\langle E\rangle + pV)/c\}$ is a four-vector whereas $\langle \boldsymbol{P}^0\rangle_0$ and $\langle E^0\rangle_0$ coincide when calculated with the methods of either A or B also follows that the method B underestimates $\langle \boldsymbol{P}\rangle$ in all frames I except the rest frame I_0.

4. A new approach to the analysis of states of the system

The preceding section demonstrates the unsatisfactory nature of the present analysis of the system under consideration from a macroscopic point of view. In order to retain the compatibility between analyses based on particle states and those based on states of the whole system, it is proposed that a radically new approach be adopted towards the latter. Whereas it has until now been usual to deal with a system in terms of the length of time it spends in each of many different states, it is now intended instead to concentrate attention upon the events which mark the transition from one state to another. These events will be characterized by incremental changes in the energy, momentum etc of the system, and will be assumed to be localized in space and time, as if caused, for example, by an instantaneous particle collision. The usefulness of this approach in dealing with Lorentz transformations between inertial frames of reference will be demonstrated in the following sections. Firstly, it is necessary to show the equivalence of the new technique and the old one.

Working in an inertial frame of reference, I, let us define a numbered set of events denoted by j, where $-\infty < j < +\infty$. Let us denote the time and place of event j by t_j and $\boldsymbol{x}_j$. We shall consider those events occurring between times t_0 and t_n inclusive, where $n \gg 1$.

Define

$$\tau \equiv t_n - t_0 \tag{4.1}$$

and

$$\pi_j \equiv (t_{j+1} - t_j)/\tau. \tag{4.2}$$

Consider an extensive quantity Q having value Q_j between events j and $j+1$. Then the main value of Q over time τ is given by

$$
\begin{aligned}
\langle Q\rangle &= \sum_{j=0}^{n-1} \pi_j Q_j \qquad (4.3)\\
&= \frac{1}{\tau}\left(\sum_{j=0}^{n-1} t_{j+1}Q_j - \sum_{j=0}^{n-1} t_j Q_j\right)\\
&= \frac{1}{\tau}\left(\sum_{j=1}^{n} t_j Q_{j-1} - \sum_{j=0}^{n-1} t_j Q_j\right)\\
&= \frac{1}{\tau}\left(\sum_{j=1}^{n-1} t_j (Q_{j-1} - Q_j) + t_n Q_{n-1} - t_0 Q_0\right)\\
&= \frac{1}{\tau}\left(-\sum_{j=1}^{n-1} t_j\, \Delta Q_j + t_n Q_{n-1} - t_0 Q_0\right) \qquad (4.4)
\end{aligned}
$$

where $\Delta Q_j \equiv Q_j - Q_{j-1}$, i.e. the increase in Q for the whole system at event j. This depends only of event j itself and not on any other event. In any other inertial frame of reference it may be obtained by the Lorentz transformation of an increment of quantity Q at a point in space–time, regardless of other changes at other points.

Equation (4.4), a summation over events, is thus equivalent to (4.3), a summation over states, though two extra terms (involving the values of Q at the start and end of the time period under consideration) are included. One of these may be reduced to zero by a suitable choice of time scale (e.g. by defining $t_0 \equiv 0$). While (4.4) appears to be of no great significance within the frame of reference in which the quantities are defined, its importance lies in the simplicity which it exhibits under the Lorentz transformation. It is of course evident that to obtain a similar expression in another frame, it is necessary to sum over a different set of events, since some of those events which lie within the specified time interval (τ) are observed from one frame will not do so when observed from another frame. Nevertheless, it is our intention to show that the sums over two such sets of events are equivalent, and hence that equation (4.4) may be transformed simply by applying the Lorentz transformation to the terms involving time and the quantity Q while still summing over the same events.

5. Validity of sums over the same set of events in different inertial frames

Consider a second inertial frame of reference I′ (quantities denoted by primed symbols), in which the original frame I moves with velocity $\boldsymbol{w}$. Define the following sets of events:

- H occurring before event O in both frames I and I′
- J occurring before event O in I and after event O in I′
- K occurring after event O in I and before event O in I′
- L occurring before event n in I and after event n in I′
- M occurring after event n in I and before event n in I′
- N occurring after event O and before event n in both frames.

It is to be assumed that the system was created at some time in the past (before event O), and that all of quantity Q which it has acquired since then is a result of such events as we are considering here. It is further assumed, since τ is taken to be large, that there are no events occurring before event O in one frame and after event n in the other. Then, from (4.4), using the symbol $\in$ to denote membership of a set of events, we have:

$$\langle Q\rangle = \frac{1}{\tau}\Big(-\sum_{j\in K,L,N} t_j\,\Delta Q_j + t_n \sum_{j\in H,J,K,L,N} \Delta Q_j - t_0 \sum_{j\in H,J} \Delta Q_j\Big)$$

$$= \frac{1}{\tau}\Big(\sum_{j\in K,L,N} (t_n - t_j)\,\Delta Q_j + \sum_{j\in H,J} (t_n - t_0)\,\Delta Q_j\Big). \tag{5.1}$$

Let us denote by $\langle Q\rangle'$ the quantity obtained by summing over the same sets of events, using the variables as they are transformed from frame I to I′. Thus:

$$\langle Q\rangle' \equiv \frac{1}{\tau'}\Big(\sum_{j\in K,L,N} (t'_n - t'_j)\,\Delta Q'_j + \sum_{j\in H,J} (t'_n - t'_0)\,\Delta Q'_j\Big). \tag{5.2}$$

However, an observer in I′ will not obtain expression (4.2) for the mean value of Q'. He will sum over a different combination of sets of events to obtain:

$$\langle Q'\rangle = \frac{1}{\tau'}\Big(\sum_{j\in J,M,N} (t'_n - t'_j)\,\Delta Q'_j + \sum_{j\in H,K} (t'_n - t'_0)\,\Delta Q'_j\Big) \tag{5.3}$$

From (5.2) and (5.3) the difference between these two quantities is given by:

$$\langle Q'\rangle - \langle Q\rangle' = \frac{1}{\tau'}\Big(\sum_{j\in J,M} (t'_n - t'_j)\,\Delta Q'_j - \sum_{j\in K,L} (t'_n - t'_j)\,\Delta Q'_j$$
$$+ \sum_{j\in K} (t'_n - t'_0)\,\Delta Q'_j - \sum_{j\in J} (t'_n - t'_0)\,\Delta Q'_j\Big)$$
$$= \frac{1}{\tau'}\Big(\sum_{j\in M} (t'_n - t'_j)\,\Delta Q'_j - \sum_{j\in L} (t'_n - t'_j)\,\Delta Q'_j$$
$$+ \sum_{j\in K} (t'_j - t'_0)\,\Delta Q'_j - \sum_{j\in J} (t'_j - t'_0)\,\Delta Q'_j\Big). \tag{5.4}$$

Examination of equation (5.4) reveals that all sums over the set N, of events which lie between O and n in both frames, have vanished. By taking the time τ' to be very much greater in magnitude than the maximum spatial dimension of the system divided by c (the velocity of light), we find that the number of events in sets J, K, L and M becomes negligibly small compared with N. Furthermore, the time intervals $(t'_n - t'_j)$, for M and L, and $(t'_j - t'_0)$, for J and K, are also negligibly small compared with τ', since events in M and L occur around event n, and events in J and K occur around event O. Equation (5.4) is thus a sum of a much smaller number of much smaller quantities as compared with equations (5.2) and (5.3). We may thus safely equate the transformed time average with the average of the transformed quantities in accordance with our intentions as expressed at the end of § 4:

$$\langle Q\rangle' = \langle Q'\rangle. \tag{5.5}$$

It is now possible, using quantities defined in frame I′, to revert to the use of a sum over states instead of a sum over events. Since, however, the order in which the terms of a summation are taken is irrelevant, we may take the events in the order in which they occur in frame I (*not* I′) and thereby ensure that each state considered in I′ begins and ends with the same pair of events as a corresponding state observed in I. One obvious consequence is that the order of such a pair may be reversed on transformation from I to

I′, and that the state it defines in I′ thus may be deemed to exist for a negative period of time. This we call a hypothetical state. It can readily be shown that despite the logical absurdity of treating such entities as if they actually existed, statistical summations over these hypothetical states produces perfectly valid results.

Thus from equations (5.5) and (5.2) we have

$$\begin{aligned}\langle Q'\rangle &= \frac{1}{\tau'}\Bigg(\sum_{j\in K,L,N}(t_n'-t_j')\,\Delta Q_j'+\sum_{j\in H,J}(t_n'-t_0')\,\Delta Q_j'\Bigg)\\ &= \frac{1}{\tau'}\Bigg(-\sum_{j\in K,L,N}t_j'\,\Delta Q_j'+t_n'\sum_{j\in H,J,K,L,N}\Delta Q_j'-t_0'\sum_{j\in H,J}\Delta Q_j'\Bigg)\\ &= \frac{1}{\tau'}\Bigg(-\sum_{j=1}^{n-1}t_j'\,\Delta Q_j'+t_n'Q_{n-1}'-t_0'Q_0'\Bigg). \qquad (5.6)\end{aligned}$$

Equation (5.6) is identical to equation (4.4) except that the symbols are primed (i.e. the quantities involved are measured in inertial frame I′). By retracing the derivation of (4.4) we obtain:

$$\langle Q'\rangle = \sum_{j=0}^{n-1}\pi_j' Q_j' \qquad (5.7)$$

where

$$\pi_j' \equiv (t_{j+1}'-t_j')/\tau'. \qquad (5.8)$$

Note that π_j' may be negative, since although $t_{j+1}\geqslant t_j$, the same relationship does not necessarily hold under Lorentz transformation. Note also that even if event j does precede event $j+1$ in frame I′, the system need not remain in the same state between those events since other events may occur between them. Thus for this reason also, the state in which the system is considered to exist between typical events j and $j+1$ may be called hypothetical (except, of course, in frame I in which the ordering of events was originally determined). Even so, the derivation of equation (5.7) holds good, and the summation over hypothetical states for the general extensive quantity Q' remains valid.

6. The occurrence under Lorentz transformation of terms involving rates of flow

It will now be shown that in the formulation of a transformed quantity in a new frame of reference, terms arise which fall into two distinct categories. The first of these depend on times measured in the earlier frame (from which the transformation is being made) and lead to expressions involving untransformed quantities measured in that frame. The others, due to the Lorentz transformation of time, depend on distances measured in the earlier frame and lead to expressions involving rates of flow of untransformed quantities measured in that frame.

In this section we will recommence using quantities measured in frame I^0 in which the system is, on average, at rest. These will be denoted by a superscript zero. Thus

$$\langle \boldsymbol{P}^0\rangle = \mathbf{0} \qquad (6.1)$$

where $\boldsymbol{P}$ denotes momentum.

Transformations will be to frame I, with a velocity in I^0 again denoted by $\boldsymbol{w}$. We define

$$\gamma \equiv \left(1 - \frac{w^2}{c^2}\right)^{-1/2}.$$

Thus, for the Lorentz transformation of time, we have

$$t_j = \gamma\left(t_j^0 + \frac{\boldsymbol{w} \cdot \boldsymbol{x}_j^0}{c^2}\right). \tag{6.2}$$

Hence, from (4.1):

$$\tau = \gamma(t_n^0 - t_0^0) + \frac{\gamma}{c^2}\boldsymbol{w} \cdot (\boldsymbol{x}_n^0 - \boldsymbol{x}_0^0)$$

By our earlier assumptions about the magnitude of the time interval τ in relation to the size of the system (which cannot be less than $(\boldsymbol{x}_n^0 - \boldsymbol{x}_0^0)$), the second term above is negligibly small, and we may thus put

$$\tau = \gamma(t_n^0 - t_0^0) = \gamma\tau^0. \tag{6.3}$$

Now, for quantity Q in the general inertial frame I, we have from (4.4), (6.2) and (6.3):

$$\begin{aligned}\langle Q\rangle &= \frac{1}{\gamma\tau^0}\left[-\sum_{j=1}^{n-1} \gamma\left(t_j^0 + \frac{\boldsymbol{w} \cdot \boldsymbol{x}_j^0}{c^2}\right)\Delta Q_j + \gamma\left(t_n^0 + \frac{\boldsymbol{w} \cdot \boldsymbol{x}_n^0}{c^2}\right)Q_{n-1} - \gamma\left(t_0^0 + \frac{\boldsymbol{w} \cdot \boldsymbol{x}_0^0}{c^2}\right)Q_0\right] \\ &= \frac{1}{\tau^0}\left(-\sum_{j=1}^{n-1} t_j^0\,\Delta Q_j + t_n^0 Q_{n-1} - t_0^0 Q_0\right) + \frac{\boldsymbol{w}}{c^2\tau^0}\left(-\sum_{j=1}^{n-1} \boldsymbol{x}_j^0\,\Delta Q_j + \boldsymbol{x}_n^0 Q_{n-1} - \boldsymbol{x}_0^0 Q_0\right).\end{aligned} \tag{6.4}$$

Each of the above terms ΔQ_j, Q_{n-1} and Q_0 is respectively derived from ΔQ_j^0, Q_{n-1}^0 and Q_0^0, together with their four-vector components, by the same linear transformation (i.e. the Lorentz transformation from frame I^0 to I). The first of the two major terms in equation (6.4) clearly resolves itself into the same linear function of $\langle Q^0\rangle$ and its components. If the second major term were zero, this would indicate that the time-averaged value for Q transformed in exactly the same way as the values of Q for each state of the system, and as the increments of Q which occur at changes of state. However, it is not always the case that the second major term is zero. In fact, as shown in the appendix, the expression

$$-\sum_{j=1}^{n-1} \boldsymbol{x}_j^0\,\Delta Q_j^0 \tag{6.4a}$$

represents the flow of quantity Q^0 through the system during time τ^0. Hence the expression

$$-\sum_{j=1}^{n-1} \boldsymbol{x}_j^0\,\Delta Q_j,$$

which occurs at the start of the second major term of equation (6.4), represents the same linear function as mentioned above applied to the flow of Q^0 and its four-vector components through the system in time τ.

Absorbing the divisor τ^0 into the term under consideration produces further simplification. Firstly the flow of Q^0 and its components is replaced by an expression for their rates of flow. Secondly, the last two expressions in the term, $\boldsymbol{x}_n^0 Q_{n-1} - \boldsymbol{x}_0^0 Q_0$, are reduced to insignificant magnitude due to our earlier assumptions about the large size of τ^0 compared with the dimensions of the system, and due also to the fact that the system is in a steady state and hence the rate of increase of Q, equal to $(Q^0_{n-1}-Q^0_0)/\tau^0$, approximates to zero.

If we now take Q to be either Lorentz invariant, or a four-vector or some higher order tensor obeying the Lorentz transformation, we may write:

$$Q_j = \psi(Q_j^0) \qquad \text{and} \qquad \Delta Q_j = \psi(\Delta Q_j^0)$$

where ψ is the linear function, dependent on $\boldsymbol{w}$, which accomplishes the Lorentz transformation from I^0 to I. It follows at once from the above arguments that the time-averaged quantity $\langle Q\rangle$ is transformed as follows:

$$\langle Q\rangle = \psi(\langle Q^0\rangle) + \frac{\boldsymbol{w}}{c^2}\cdot\psi(\langle\dot{\boldsymbol{Q}}^0\rangle) \tag{6.5}$$

where the three-vector $\langle\dot{\boldsymbol{Q}}^0\rangle$ denotes the time-averaged rate of flow of quantity Q^0 through the system as measured in frame I^0. It is the presence of the second term in equation (6.5) which shows the validity of the statements made in the first paragraph of this section. This equation may be rewritten in tensorial notation for convenience in handling higher order quantities. Thus:

$$\langle Q\rangle_{\mu\ldots} = \psi^{\nu\ldots}_{\mu\ldots}\langle Q^0\rangle_{\nu\ldots} + \frac{1}{c^2}\psi^{\nu\ldots}_{\mu\ldots}\langle\dot{Q}^0\rangle_{k,\nu\ldots}w^k \tag{6.6}$$

where μ, $\nu = 1\to 4$, $k = 1\to 3$; $\nu\ldots$ represents a tensor of indefinite order.

7. Application of the above theories in the case of energy and momentum

The ideas embodied in equation (6.5) can now be applied to specific instances. Let us therefore consider the energy and momentum of a system confined inside a closed vessel, with pressure p and volume V^0 (as measured in the inertial frame I^0). Firstly, note that the rate of flow of energy through the stationary system is zero; this of course follows at once from the fact that no work is being done at any part of the system's surface, as well as from its having zero momentum. However, momentum does flow through the system due to the forces constraining it within the closed vessel. Familiar arguments lead to the result that for any chosen direction, momentum in that direction flows through the system in that same direction with magnitude pV^0. Note that if momentum directed in one way flowed in a different direction, this would indicate the presence of lateral stresses within the system; while not physically impossible, this would not apply in the types of system being considered here (e.g. a gas in equilibrium). The 3×4 tensor for the rate of flow of energy-momentum $\langle U^0\rangle_\mu$ through the system is therefore:

$$\langle U^0\rangle_{k,\nu} = \begin{pmatrix} pV^0, & 0, & 0, & 0 \\ 0, & pV^0, & 0, & 0 \\ 0, & 0, & pV^0, & 0 \end{pmatrix}. \tag{7.1}$$

In the case of energy-momentum, the 4×4 tensor ψ^ν_μ which transforms from I^0 to I takes the following well known form (it is here assumed that we have chosen Cartesian coordinates such that $\mathbf{w}$ is parallel to the first spatial coordinate):

$$\psi_\mu^\nu = \begin{pmatrix} \gamma, & 0, & 0, & \gamma w/c^2 \\ 0, & 1, & 0, & 0, \\ 0, & 0, & 1, & 0 \\ \gamma w, & 0, & 0, & \gamma \end{pmatrix}. \tag{7.2}$$

We also have for energy-momentum $\langle U^0\rangle_\mu$ and velocity w^k

$$\langle U^0\rangle_\nu \equiv \begin{pmatrix} \langle P_x^0\rangle \\ \langle P_y^0\rangle \\ \langle P_z^0\rangle \\ \langle E^0\rangle \end{pmatrix} \tag{7.3}$$

and

$$w^k \equiv \begin{pmatrix} w \\ 0 \\ 0 \end{pmatrix}. \tag{7.4}$$

Replacing Q by U and substituting (7.1)–(7.4) in (6.6) gives:

$$\langle U\rangle_\mu \equiv \begin{pmatrix} \langle P_x\rangle \\ \langle P_y\rangle \\ \langle P_z\rangle \\ \langle E\rangle \end{pmatrix} = \begin{pmatrix} \gamma\langle P_x^0\rangle + \gamma(w/c^2)\langle E^0\rangle \\ \langle P_y^0\rangle \\ \langle P_z^0\rangle \\ \gamma w\langle P_x^0\rangle + \gamma\langle E^0\rangle \end{pmatrix} + \frac{1}{c^2}\begin{pmatrix} \gamma w p V^0 \\ 0 \\ 0 \\ \gamma w^2 p V^0 \end{pmatrix}.$$

Therefore

$$\langle U\rangle_\mu \equiv \begin{pmatrix} \langle P_x\rangle \\ \langle P_y\rangle \\ \langle P_z\rangle \\ \langle E\rangle \end{pmatrix} = \begin{pmatrix} \gamma[\langle P_x^0\rangle + (w/c^2)(\langle E^0\rangle + pV^0)] \\ \langle P_y^0\rangle \\ \langle P_z^0\rangle \\ \gamma[\langle E^0\rangle + (w^2/c^2)pV^0 + w\langle P_x^0\rangle] \end{pmatrix} = \begin{pmatrix} \gamma(w/c^2)(\langle E^0\rangle + pV^0) \\ 0 \\ 0 \\ \gamma(\langle E^0\rangle + (w^2/c^2)pV^0) \end{pmatrix} \tag{7.5}$$

since the momentum is zero in I^0.

Equation (7.5) incorporates the well known expressions for the Lorentz transformation of the energy and momentum of a confined system from the inertial frame in which it is at rest to another inertial frame.

8. Discussion

In the preceding sections we have seen contrasted two different approaches to the statistical problems presented by a compressed gas in the special theory of relativity. The first of these (§ 2 and § 3) is based on earlier theories, and by the use of the concept of a 'state indicator' attached to the system, enables the difficulties found in previous work to be overcome. The second (§§ 4–7) adopts a radically new approach and enables similar results to be obtained without explicit consideration of particle states, and without recourse to the artificial 'state indicator'. The equally artificial 'hypothetical states' of the system are shown to be mathematically valid, yielding no unexpected anomalies in the cases considered, even though they do not correspond to the actual system states. Both theories can of course be extended to cover Lorentz transformations between any two inertial frames, since the transformation between I^0 and a general frame is known.

The second of them can particularly easily provide transformation formulae for other quantities than those considered here, by means of equations (6.5) and (6.6). These can be rendered into a covariant form by dividing throughout by the volume of the system (V on the left, V^0 on the right, producing an extra factor γ, since $V^0 = \gamma V$). It can then be seen that the first term on the right-hand side is the product of the fourth (time-like) components of two four-vectors (or a vector and a higher order tensor), and the second term is the product of the three space-like components of the same entities. In the case of energy-momentum, this will bring us back to the transformation of the energy density/momentum density/pressure tensor, commonly called $T_{\mu\nu}$. It should not be thought, however, that the calculations in §§ 4–7 above apply only to systems of compressed gas. They can equally, for example, be used to consider the flow of particles through an open ended tube. The mean number of particles in the tube in any frame I can be calculated from the mean number and rate of flow in I^0 by use of equation (6.5), and the validity of the result can easily be checked from well known results concerning the transformation of volume and particle density. This is of course a trivial result, but it serves to show the generality of the theory.

An important part of this theory is that it involves states of the whole system, and not of any particles which may comprise it. The fact that the events at which the system changes state may well be particle collisions does not affect this point. While the criticisms of earlier works in §§ 2 and 3 are valid as far as the methods used in those works are concerned, we do not suggest that the approach used in those works (based on states of the system) is in itself incorrect. Rather we suggest that both approaches are inherently acceptable, though we realize that there are drawbacks in the application of each. In the one case (§§ 4–7) there are 'hypothetical states' existing for negative periods of time; in the other (§§ 2–5 of I, §§ 2 and 3 of this paper) there is the 'state indicator' which, being based on the use of a preferred frame of reference, is a manifestly non-covariant concept. The 'kinematic weight factors' used in this latter scheme bear a strong resemblance to the Lorentz-transformed expressions for the time spent in the 'hypothetical states' of the former. While, however, the mathematical backing for the use of 'hypothetical states' is readily available (being derived from consideration of the events at which the system changes shape), the 'kinematic weight factors' depend for their existence on the use of the non-covariant 'state indicator'. Which theory is to be used in any given situation thus requires careful consideration.

Appendix. Discussion of equation (6.4*a*)

The problem of quantifying the flow of the general extensive quantity Q^0 through a system should be approached by considering first of all what we mean by the flow of such a quantity. Let us observe that Q^0 can be localized in the sense that we can say that a certain particle, or body or region of space contains a certain measure of this quantity (say Q^0_α) at some instant in time, while another particle etc may contain another measure (say Q^0_β) at the same or a different instant. It is then not unreasonable to say (indeed it may even be taken as a definition) that the flow of Q^0 is the sum over all particles, bodies etc of the product of the measure of Q^0 attached to a particle α and the distance moved by that particle while it remains in the system, i.e.

$$\text{flow of } Q^0 = \sum_\alpha \Delta \boldsymbol{x}^0_\alpha \, . \, \Delta Q^0_\alpha. \tag{A.1}$$

Dividing such an expression by the increment of time in which the various motions take place gives a rate of flow of Q^0 equal to the sum of the products of particle velocities and corresponding measures of Q^0. This readily yields familiar expressions for various rates of flow (e.g. momentum, the rate of flow of mass, is given by the sum of the products of particle velocities and particle masses).

The difficulty arises in the present case because there are no particles or bodies envisaged in our latest model of a system. Let us therefore rearrange equation (A.1) so as to eliminate the explicit reference to particles etc, α, and replace it by reference to events in space and time. Thus let us assign to each particle α a location $\boldsymbol{x}^0_\alpha$ where it is to be found at the start of the period of time in which it is in the system; clearly at the end of this period each particle will be at $\boldsymbol{x}^0_\alpha + \Delta \boldsymbol{x}^0_\alpha$. Equation (A.1) then becomes

$$\text{flow of } Q^0 = \sum_\alpha (\boldsymbol{x}^0_\alpha + \Delta \boldsymbol{x}^0_\alpha) \, \Delta Q^0_\alpha + \sum_\alpha (-\boldsymbol{x}^0_\alpha) \, \Delta Q^0_\alpha. \tag{A.2}$$

Observe that the first of the two sums above corresponds to events where particles cease to be considered as contributing to the flow (i.e. events at the ends of the time intervals); on the other hand, the second sum corresponds to events where a particle begins to contribute (i.e. events at the beginnings of the time intervals). Together they form a sum over all events where positive or negative increments of Q^0 are added to the system. Let us again rearrange the expressions, this time as sums over the sets of events $\{J+\}$ (positive increments of Q^0, where particles enter the system), and $\{J-\}$ (negative increments of Q^0, where particles leave the system). For events j in $\{J+\}$, the increments ΔQ^0_j are clearly equal to the corresponding increments ΔQ^0_α, whereas for events in $\{J-\}$ they are equal to $-\Delta Q^0_\alpha$. Similarly the locations $\boldsymbol{x}^0_j$ of events in $\{J+\}$ correspond to each $\boldsymbol{x}^0_\alpha$ while those in $\{J-\}$ correspond to each $(\boldsymbol{x}^0_\alpha + \Delta \boldsymbol{x}^0_\alpha)$. Equation (A.2) immediately becomes

$$\begin{aligned} \text{flow of } Q^0 &= \sum_{j \in \{J-\}} \boldsymbol{x}^0_j (-\Delta Q^0_j) + \sum_{j \in \{J+\}} (-\boldsymbol{x}^0_j) \, \Delta Q^0_j \\ &= - \sum_{j \in \{J- \cup J+\}} \boldsymbol{x}^0_j \, \Delta Q^0_j. \end{aligned} \tag{A.3}$$

It can therefore be seen that so long as the set of events over which the sum is taken consists of events where particles start and finish contributing their increments ΔQ^0_α to the system, equation (A.3) is as valid as our original definition of the flow of Q^0, as expressed in equation (A.1). Furthermore, though our model of the system does not deal with particles explicitly we certainly do not claim that particles do not exist or that an extensive quantity Q^0 is not quantizable into discrete increments. Thus the only consideration which could invalidate the use of equation (A.3) in our statistical model would be an overall increase or decrease of Q^0 over the long periods of time τ^0 which we consider, indicating that there was an imbalance of positive and negative increments of Q^0. This, however, would contradict our basic assumption concerning the system being in a ready state, and therefore we can assert that within the limitations of our model, the rate of flow of a general extensive quantity Q^0 is given by

$$-\sum_j \boldsymbol{x}_j^0 \, \Delta Q_j^0$$

where the sum is taken over all events where an increment of Q^0, ΔQ_j^0 (positive or negative), is added to the system.

Acknowledgments

We are grateful to Professor K Kraus, Würzburg, for collaboration and in particular for the inequivalence proof of § 3.

References

Anderson J L and Witting H R 1974 *Physica* **74** 466
Grøn Ø 1973 *Nuovo Cim.* B **17** 141–65
Habeger C C 1972 *Ann. Phys, NY* **72** 1–28
Kraus K and Landsberg P T 1974 *J. Phys. A: Math., Nucl. Gen.* **7** 2130–43
Landsberg P T 1970 *A Critical Review of Thermodynamics* eds E B Stuart, B Gal-Or and A J Brainard (Baltimore: Mono Book Corporation) pp 253–74
Landsberg P T and Johns K A 1970a *J. Phys. A: Gen. Phys.* **3** 113–20
—— 1970b *Proc. Int. Conf. on Thermodynamics, Cardiff 1970* (London: Butterworths) pp 487–92
Szamosi G 1973 *Found. Phys.* **3** 241–6