On the Nature of Theories of Irreversible Processes

TA-YOU WU

*Statistical Physics Laboratory, Department of Physics,
State University of New York, at Buffalo, New York*

Abstract

The purpose of the present article is to give an elementary discussion of the nature of the theories used to describe the observed irreversible processes in gases to the equilibrium state, with special regard to the explicit mathematical time-reversal invariance or non-invariance of the theories.

1. Introduction

1.1. *Macroscopic Irreversibility and Empirical Laws*

We are concerned with the non-equilibrium behavior of a gas. In order to be definite, let us state the nature of our observations. Consider a gas in an enclosure, which at a certain instant is in a non-equilibrium state. Observation shows that within a short time the gas approaches a steady state in which the molecular velocities are distributed according to Maxwell's law and the spatial distribution is uniform (in the case of no external field). Experience also shows that the approach to equilibrium, except for small fluctuations, is monotonic, the fluctuations being the smaller the greater is the average density of the gas. Other familiar irreversible processes are the mixing of gases by diffusion, the production of heat in viscous flows and the conduction of heat.

The above description of these processes as irreversible is a macroscopic view and is expressed in terms of concepts, or macroscopic variables, such as the density, temperature, pressure, etc. The mathematical expressions of the irreversibility are the familiar Fourier equations of heat conduction, the equation of diffusion, and the Navier-Stokes equation for viscous flow. All these equations are not invariant upon time reversal, i.e., are irreversible in time. This irreversibility is not the consequence of the basic dynamical laws governing the motion of the molecules; it is introduced into the equations on an empirical basis. We shall call these theories 'phenomenological' theories, to distinguish them from others which we attempt to construct on the basis of more fundamental theories.

21   325

1.2. *Thermodynamics*

At the macroscopic level, classical thermodynamics deals with the properties of matter in the equilibrium state. The laws of thermodynamics are based on empirical experience in dealing with macroscopic phenomena. Even the axiomatic form of Caratheodory for the Second Law still depends on the validity of an empirical fact. The methods of thermodynamics are general in the sense that they do not depend on any specific or detailed knowledge of the system at the atomic level. For this reason, the thermodynamic method is powerful and its range of application wide. But it is for the same reason that it has its limitations. Thus thermodynamics can give no information on such a simple question as the equation of state of a given gas.

In thermodynamics, the concept of macroscopically irreversible processes plays a very basic role. The Second Law definitely determines the direction of the natural processes in a system not in thermodynamic equilibrium, namely, the direction of increasing entropy. This irreversibility is not explained by the Second Law; rather, it is a part of the empirical experience that has gone into the formulation of the Second Law itself (in the induction and generalization process).

Modern non-equilibrium (or irreversible) thermodynamics has been developed to treat certain problems, in particular, processes in systems having small deviations from the equilibrium state. The conditions for the stability of an equilibrium state according to the Second Law determine the irreversible returns to the equilibrium state. A most important development is Onsager's (1931) theory of irreversible thermodynamics, which is beyond the scope of the present article. It is, however, so important in the subject of irreversible processes that a brief reference will be made below.

Consider a system in which there are temperature and density gradients. The heat flux J and the mass flux N can be expressed by equations of the form

$$J = a_{11} \nabla T + a_{12} \nabla n$$

$$N = a_{21} \nabla T + a_{22} \nabla n$$

where a_{11}, a_{22} are, up to proportional factors, the coefficients of heat conduction and diffusion, $a_{21} \nabla T$ gives the thermal diffusion and $a_{12} \nabla n$ the heat flow due to diffusion. These equations are the generalizations of the heat conduction and diffusion equations and are themselves parts of an irreversible theory. They can be regarded as phenomenological equations, or they can be derived from such a theory as Boltzmann's transport equation (Section 3.1), which itself is an irreversible theory.

In this particular case, Onsager's theory gives the following reciprocity relation (when all quantities are expressed in units of appropriate dimensions)

$$a_{12} = a_{21}$$

Such reciprocity relations, however, are not empirical in origin but can be proved from general principles such as the fundamental reversibility of microscopic processes.

Thus, important as the reciprocity relations are, they do not bear on the question of the basis of irreversible theories or on the origin of the observed macroscopic irreversibility, which is the subject of the present article.

1.3. *Statistical–Kinetic Theories*

In the kinetic theory of gases, one starts with a system of molecules. One is, however, not interested in a detailed knowledge of the individual molecules, but only in some macroscopic properties of the gas. These macroscopic properties must then be related to some average properties of the molecules. It is in dealing with extremely large numbers (Avogadro's number $= 6 \times 10^{23}$) that the probability concepts and statistical methods enter into the theory.

For systems in equilibrium, theories have been developed in which some basic postulates form the foundation. These theories have been called statistical mechanics. For systems not in equilibrium, equilibrium thermodynamics and equilibrium statistical mechanics are not sufficient. The aim is to formulate theories for describing the observed irreversible processes in the macroscopic view.

There have been many different theories of irreversible processes. They can be grouped into two categories. In one, the theories are based on some assumptions of a probability nature. To this category belong the Boltzmann transport theory, the Fokker–Planck theory and the so-called Master Equation (theory). In these theories, the basic equation is time-reversal non-invariant, or as one may say, has a time arrow. In the other category are the theory of Boltzmann and Gibbs, and the many recent theories of Bogoliubov, Born, Green, Kirkwood and others. All these theories start from the Liouville equation, which is the consequence of classical dynamical laws and is time-reversal invariant.

In the following sections we shall discuss the nature of these theories.

1.4. *Quantum Theory*

In quantum mechanics there is an analogue to each of the basic classical features. The classical equations of motion and the Schrödinger equation are both time-reversal invariant. To the Liouville equation (Section 2.2) for the density in phase space, there is an exact quantum analogue for the density matrix, and both are time-reversal invariant. To the classical Master Equation (Section 3.3) for the probability, there is an exact quantum analogue, called the Pauli equation, for the diagonal elements of the density matrix. Both equations are irreversible in time, the classical one on account of the Smoluchowski hypothesis for Markovian processes, the quantum one because of the random phase hypothesis.

It is generally believed that the observed macroscopic irreversibility is

of classical statistical origin and is not a quantum effect. For this reason, we shall confine our discussions in the present article to the classical theories.

2. *Statistical–Kinetic Theory of Boltzmann and Gibbs*†

2.1. *Poincaré Theorem* (1890)

In the statistical–kinetic theory, a gas is a system of molecules whose motions are governed by the dynamical laws. Let $\mathscr{H}(q,p)$ be the Hamiltonian of the system, where q,p stands for $q_1, \ldots, q_N, p_1, \ldots, p_N$. The equations of motion are

$$\dot{q}_k = \frac{\partial \mathscr{H}}{\partial p_k}, \qquad -\dot{p}_k = \frac{\partial \mathscr{H}}{\partial q_k}, \qquad k = 1, 2, \ldots, N \qquad (2.1.1)$$

In the $6N$-dimensional Γ phase space of q,p, the dynamical state of the system is represented by a point $P(q,p)$. If the state at a certain instant t_0 is given, say $P_0(q^0,p^0)$, then the whole history of the system is represented by the trajectory through P_0. As equations (2.1.1) are first-order differential equations in time, there is one, and only one, trajectory through any point which is uniquely determined by the equations (2.1.1). Therefore, a trajectory never crosses itself, nor another, at any time. Hence, in the course of time, the phase point P goes all over the accessible portion of the Γ-space, for example, the hypersurface corresponding to the relation: total energy = constant (or, an energy shell of thickness ΔE).

From Poincaré's theorem, the trajectory of the phase point of a gas contained in a finite volume will pass, in the course of time, arbitrarily close to the point P_0 itself, without going through it. The time of this quasi-cycle depends on how close one wants it to approach P_0, and is extremely long for large N. The length of this quasi-cycle has been estimated by Boltzmann.† Thus for a gas containing 10^{18} molecules in 1 cc with mean velocity 5×10^4 cm/sec, the time required for all the molecules to come within a range of 10^{-7} cm and 10^2 cm/sec in each of the coordinate and velocity components respective to the initial P_0 is greater than $10^{10^{19}}$ years (Chandrasekhar, 1943). This theorem is of basic importance in our understanding of the 'irreversibility' of the approach of a gas toward equilibrium.

2.2. *Liouville Equation*

Let $\rho(q,p,t)\,dq\,dp$, where $dq\,dp$ stands for $dq_1, \ldots, fq_N, dp_1, \ldots, dp_N$, be the probability of the phase point $P(q,p)$ being the volume element $dq\,dp$, so that

$$\int \cdots \int \rho \, dq \, dp = 1 \text{ for all } t \qquad (2.2.1)$$

† L. Boltzmann's work appeared in the Journal *Sitzungsberichte der Akademie der Wissenschaften in Wien*, during the years 1860 to 1890. A good deal of the discussions can be found in his *Vorlesungen über Gastheorie*. Leipzig (1896–98); and its English translation in *Lectures on Gas Theory*. University of California Press, Berkeley (1964). See also, Gibbs, J. W. (1902). *Elementary Principles in Statistical Mechanics*. This work is included in his *Collected Works*, Vol. II. Yale University Press (1948).

the integral being taken over the accessible volume. This convention will always be used in the following discussion. It follows from the equations (2.1.1) that any volume in the Γ space remains unchanged in time if we follow the points in that volume along their trajectories, although the shape of the volume will change. From this and the conservation of probability in that volume, the Liouville equation follows (Liouville, 1838):

$$\frac{\partial \rho}{\partial t} = -\sum_i \left(\frac{\partial \rho}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial \rho}{\partial p_i} \frac{\partial \mathcal{H}}{\partial q_i} \right) \tag{2.2.2}$$

An important property of the Liouville equation is that it is time-reversal invariant, i.e., upon time reversal

$$t \to -t = \tau, \qquad q \to q, \qquad p \to -p, \qquad \rho(q,p,t) \to \bar{\rho}(q,-p,\tau) \tag{2.2.3}$$

equation (2.2.2) becomes

$$\frac{\partial \bar{\rho}}{\partial \tau} = -\sum_i \left(\frac{\partial \bar{\rho}}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial \bar{\rho}}{\partial p_i} \frac{\partial \mathcal{H}}{\partial q_i} \right)$$

which has the same form as (2.2.2).

From (2.2.2), it follows that for any function $F(\rho)$ of ρ,

$$\frac{d}{dt} \int \cdots \int F(\rho) \, dq \, dp = 0 \tag{2.2.4}$$

If we define the function $H(t)$

$$H(t) \equiv \int \cdots \int \rho \ln \rho \, dq \, dp \tag{2.2.5}$$

then from (2.2.4), it follows that

$$\frac{dH(t)}{dt} = 0 \tag{2.2.6}$$

If we look at H in (2.2.5) as a functional of $\rho(q,p)$, the requirement that H be stationary subject to (2.2.1) as an auxiliary condition leads to

$$\rho = \text{constant} \tag{2.2.7}$$

From the second variation and $\rho > 0$, we find:

$$H(t \,|\, \rho) \text{ is minimum if } \rho \text{ is a constant, independent of } q, p \tag{2.2.8}$$

In this connection, let us prove the following theorem. Let a function $P = P(q,p,t)$, $P \geqslant 0$, satisfy the condition

$$\int \cdots \int P(q,p,t) \, dq \, dp = 1 \tag{2.2.9}$$

Also let $\rho(q,p,t)$ be any function satisfying

$$\rho \geqslant 0, \qquad \int \cdots \int \rho \, dq \, dp = 1$$

Define the functional $\bar{H}$ of $P(q,p,t)$ by

$$\bar{H}(t|P) \equiv \int \cdots \int \rho(q,p,t) \ln P \, dq \, dp \qquad (2.2.10)$$

It can readily be shown, for example by the variational method, that

$\bar{H}(t|P)$ has a maximum value when P equals ρ'   (2.2.11)

Hence we obtain

$$\int \cdots \int \rho \ln \rho \, dq \, dp \geqslant \int \cdots \int \rho \ln P \, dq \, dp \qquad (2.2.12)$$

the equality sign holding when $P = \rho$. Note that this theorem (2.2.11), or (2.2.12), is purely a mathematical result.

2.3. Theory of Boltzmann†

   To obtain the essential statistical features of the problem, we shall consider a system of non-interacting and therefore independent particles. In this case it is sufficient to consider the 6-dimensional μ phase space of one particle. Imagine the μ space as being divided into a large number of cells each of volume ω. ω must be small on the macroscopic scale but must not be too small, so that each cell still contains a number of particles. Let n_1, n_2, ... be the number of molecules in cells 1, 2, ... The number of permutations of N molecules in the above distribution is

$$G_{n_1, n_2, \ldots} = \frac{N!}{n_1! \, n_2! \, \ldots}, \qquad \sum n_i = N$$

The weight of each distribution is $(\omega/V)^N$, where V is the volume of the accessible μ space. The distribution that can be realized in the largest number of ways is found by maximizing G subject to the conditions $\sum n_i = N$ and $\sum n_i \epsilon_i = E$, where ϵ_i is the energy of a molecule when its phase lies in the cell ω_i, and E is the total (constant) energy of the gas. The calculations are familiar, and we shall only state the qualitative result here. It is found that G has a very strong and sharp maximum, it being the sharper the larger N is. This result means that of all the possible permutations and distributions of the N molecules, an overwhelmingly large number corresponds to the few distributions near the maximum of G. The distribution giving the maximum G is the most probable state. Boltzmann identified this most probable state with the equilibrium state.

   The probability of the distribution n_1, n_2, n_3, ... is now

$$W_{n_1, n_2, \ldots} = G_{n_1, n_2, \ldots} (\omega/V)^N$$

and

$$\sum W_{n_1, n_2, \ldots} = 1$$

the summation being taken over all partitions n_1, n_2, ... of N.

---

† See footnote on p. 328.

Let us translate the above result to the Γ space. In the Γ space, ω^N is the volume Ω of a cell and V^N is the volume of the Γ space. The sharp maximum for $G_{n_1, n_2, \ldots}$ now means that an overwhelmingly large portion of the Γ space corresponds to the most probable state which is identified with the equilibrium state.

Consider now a gas which at an instant t_0 is in a non-equilibrium state represented by a point P_0. According to the discussions in Section 2.1, the phase point will move endlessly in the Γ space. Since an overwhelmingly large proportion of the Γ space corresponds to the equilibrium state, the phase point will pass from P_0 in a time of the order of the relaxation time into this large portion of the Γ space, and the system will approach rapidly, and will mostly be found in, the equilibrium state.

The question as to the non-observation of the recurrence of the initial state P_0 is answered by noting the extremely long quasi-period of the Poincaré quasi-cycle. In the Γ space, there are of course many points Q which correspond to highly non-equilibrium states. However, starting from any point P_0, it also takes a very long time in general to come very close to one of them. Compared to the points corresponding to the equilibrium state, these non-equilibrium points have measure zero. Rigorously speaking, there is no irreversibility. But the reverse of the approach to equilibrium is so improbable that for practical purposes it can be excluded.

As the phase point P moves endlessly all over the Γ-space, it is plausible to assume that on average it spends an equal amount of time in volume elements of equal size at different points in the Γ-space, i.e., the probability of the phase point P being found anywhere will be the same. This means that ρ approaches a constant value over the Γ-space. As an overwhelmingly large proportion of the Γ-space corresponds to the equilibrium state, this suggests that the equilibrium state corresponds to a uniform ρ.

In Section 2.2, the $H(t)$ function defined in equation (2.2.5) is minimum for a constant value of ρ. Thus one would have a theory for an irreversible approach to equilibrium in a gas if one could show that $H(t)$ keeps on decreasing until it reaches its minimum at which ρ is uniform, i.e., if

$$\frac{dH}{dt} < 0 \tag{2.3.1}$$

This hypothesis is known as the H theorem of Boltzmann (1872, 1875).

We have seen in equation (2.2.6), however, that as a consequence of the Liouville equation, $dH/dt = 0$. Thus the H theorem cannot be established on the basis of the laws of dynamics alone. There had been the criticism of Loschmidt (1877) that if $H(t)$ decreases with time for one state of motion, then it will have to increase with time for another state in which all the molecular velocities are just opposite to those of the first state. There had also been the criticism of Zermelo (1896) that since the state in the course of time returns arbitrarily close to any initial state according to Poincaré's theorem, H could not be monotonically decreasing at all times.

To sum up, we may state the following. In the probability sense, it is

possible to understand the macroscopically irreversible approach of a gas, from any initial state, to the equilibrium state. The H theorem, however, cannot be established on dynamical laws alone and attempts to maintain it must be made on probability grounds, with additional assumptions outside of dynamical laws.

2.4. *Theory of Gibbs—Microcanonical Ensemble* (Tolman, 1935)†

Gibbs introduced the concept of ensembles. An ensemble is a very large number of systems which all have the same macroscopic variables as, but different dynamical states (q,p) to, the system under consideration. The ensemble is represented by a distribution function $\rho(q,p,t)$ in the Γ-space which is now normalized to N instead of unity in equation (2.2.1). All the results of Section 2.2 hold.

Consider those points that lie within a volume element $\Delta = dq\,dp$, namely, $\rho\Delta$. Each point will move in accordance with the equations of motion (2.1.1) as described in Section 2.1, and the points $\rho\Delta$ will move as an incompressible fluid, keeping volume Δ but changing shape. In the course of time these points will deform into a thin filament over the accessible part of the Γ-space. Similarly, for other systems lying outside Δ. The result is that, from any initial distribution $\rho(q,p)$, the ensemble will in the course of time spread out 'uniformly' over the Γ-space.

In Boltzmann's theory, it is assumed that the equilibrium properties of a gas are given by a long time average (i.e., long compared with all the relaxation times of the system). In Gibbs's theory, the fundamental hypothesis is made that the time average over one system is equal to the average taken over the ensemble. As the phase points of an ensemble extend, in time, over the Γ-space 'uniformly', the ensemble average (average over the Γ-space) gives the equilibrium value.

These qualitative discussions make plausible, in the probability sense, the 'irreversible' approach of a system from any non-equilibrium state (for example, all systems initially lying within a volume Δ) to the equilibrium state represented by a 'uniform' distribution in the Γ-space. The words 'uniform' and 'uniformly' are to be taken in the coarse-grained sense defined below.

The concept of coarse-graining is introduced on the following considerations. In the macroscopic observations made on a gas, one does not know the positions and momenta of the molecules except to within ranges Δq and Δp. The coarse-grained density $\bar{\rho}$ is defined as the average value of the 'fine-grained' ρ over a volume $\Delta = \Delta q\,\Delta p$ of size corresponding to the limits of accuracy of experimental observation, i.e.,

$$\bar{\rho}(q,p) = \frac{1}{\Delta}\int\cdots\int_{\Delta} dq\,dp\,\rho \qquad (2.4.1)$$

† See also, footnote on p. 343.

$\bar{\bar{\rho}}$ is a function of phase (q,p), but is constant over the volume $\varDelta$. With this definition, we obtained for the coarse-grained ensemble average of a function $F(\rho)$

$$\bar{F} = \int \cdots \int \bar{\bar{\rho}} F(\bar{\bar{\rho}}) \, dq \, dp = \int \cdots \int \frac{1}{\varDelta_{qp}} \int_{\varDelta_{qp}} \cdots \int \rho(q',p') \, dq' \, dp' \, F(\bar{\bar{\rho}}(q,p)) \, dq \, dp$$

$$= \int \cdots \int \rho F(\bar{\bar{\rho}}) \, dq \, dp \tag{2.4.2}$$

Let us now come back to the ideas of Gibbs on the spreading of the phase points in $\varDelta$ over the $\varGamma$-space. The foregoing qualitative discussions make it plausible that $\rho(q,p)$ will approach a uniform coarse-grained $\bar{\bar{\rho}}$. Now, we may define a coarse-grained H function $\bar{H}(t)$ by (2.4.2)

$$\bar{H}(t) = \int \cdots \int \bar{\bar{\rho}} \ln \bar{\bar{\rho}} \, dq \, dp$$

$$= \int \cdots \int \rho \ln \bar{\bar{\rho}} \, dq \, dp \tag{2.4.2a}$$

Using the theorem in equation (2.2.12), we immediately obtain

$$\int \cdots \int \rho \ln \rho \, dq \, dp \geqslant \int \cdots \int \rho \ln \bar{\bar{\rho}} \, dq \, dp \tag{2.4.3}$$

or,

$$H(t) \geqslant \bar{H}(t) \tag{2.4.3a}$$

i.e., the fine-grained H is always greater than the coarse-grained $\bar{H}$. It is to be noted that the relation (2.4.3) or (2.4.3a) is a consequence of the definition (2.4.1) and the mathematical inequality (2.2.12), and has nothing to do with the time variation of the ensemble or the system in question.

We have seen in equation (2.2.6) that the fine-grained H remains constant in time. The coarse-grained $\bar{H}$ as defined by (2.4.1) is not governed by the Liouville equation, and one may ask whether $\bar{H}$ decreases monotonically in time, i.e., whether

$$\frac{d\bar{H}}{dt} < 0 \tag{2.4.4}$$

If the coarse-grained $\bar{H}$ satisfies the H theorem (2.4.4), then (2.4.4) is the mathematical expression of the qualitative ideas of Gibbs on the approach of a gas from any arbitrary state to equilibrium.

To express Gibbs's ideas in terms of the coarse-grained density, let us prepare an ensemble such that at an instant t_1 (Tolman, 1935; ter Haar, 1954, 1955),

$$\rho_1 = \text{constant} = \bar{\bar{\rho}}_1 \text{ inside a volume } \varDelta$$

$$= 0 \text{ elsewhere} \tag{2.4.5}$$

At a later time t_2, the phase points will have spread so that in general

$$\rho_2 \neq \bar{\bar{\rho}}_2 \tag{2.4.6}$$

Then

$$\bar{H}(t_2) - \bar{H}(t_1) = \int \cdots \int \bar{\rho}_2 \ln \bar{\rho}_2 \, dq \, dp - \int \cdots \int \bar{\rho}_1 \ln \bar{\rho}_1 \, dq \, dp$$

by (2.4.2) and (2.4.5),

$$= \int \cdots \int \rho_2 \ln \bar{\rho}_2 \, dq \, dp - \int \cdots \int \rho_1 \ln \rho_1 \, dq \, dp$$

by (2.2.6),

$$= \int \cdots \int \rho_2 \ln \bar{\rho}_2 \, dq \, dp - \int \cdots \int \rho_2 \ln \rho_2 \, dq \, dp$$

and by (2.4.3)

$$< 0 \tag{2.4.7}$$

Thus $\bar{H}$ does decrease from the initial condition (2.4.5) at $t = t_1$ to a smaller value at any later time $t_2 > t_1$.

The above proof depends entirely on the special initial condition $\rho_1 = \bar{\rho}_1$ in (2.4.5). Without it, it would not have been possible to proceed beyond the first line in (2.4.7). What has been, and can be, proved is that the initial non-equilibrium state (2.4.5) with $\rho_1 = \bar{\rho}_1$ in Δ does evolve with a decrease in $\bar{H}$, and this is the qualitative discussions of Gibbs's ideas given before. There is, however, no proof that $\bar{H}(t_3) - \bar{H}(t_2) < 0$ for a later time $t_3 > t_2$, i.e., there is no proof that $\bar{H}$ decreases monotonically, in contrast to the arguments of Tolman that $\bar{H}$ continues to decrease (Tolman, 1935). It is true that $\bar{H}$, unlike H, is not bound by any law to remain constant, but the H theorem for the coarse-grained $\bar{H}$ cannot be proved without some additional assumption concerning the time behavior of the coarse-grained density $\bar{\rho}$. Through the years, the H theorem has been the subject of discussions, by Boltzmann, Loschmidt, Zermelo and the Ehrenfests, and it became clear that the theorem must be interpreted on a probability basis. It will be shown in Section 3.3 that an assumption on the time relation of the state probabilities does introduce a time arrow and lead to an H theorem.

To summarize, we may conclude as follows. On the basis of the Liouville equation, although no mathematically irreversible law (such as the H-theorem) can be derived, it is still possible to 'understand' the irreversible approach of a gas to equilibrium in the probability sense. The basic reason is because of the large number of permutations of molecules that all correspond to the equilibrium (the most probable) state and the long quasi-cycle, by virtue of the large number of molecules in a macroscopic amount of gas.

3. Theories with Explicit Time-Reversal Non-Invariance

We shall next consider a few theories of the other category mentioned in Section 1.3, namely, those that are by construction, by making assumptions of some sort, explicitly time-reversal non-invariant. Such theories do

not claim to be the consequences of the fundamental dynamical laws and are therefore not to be criticized on that basis. The criteria will be their plausibility, simplicity and ultimately their success in applications to actual problems.

3.1. *Boltzmann's Theory*

The earliest theory of transport processes is that of Boltzmann. Instead of dealing with the N molecules of a gas in the Γ-space, it is assumed that a theory for describing the gas can be formulated in terms of the one-particle distribution function $f(r,v,t)$ in the μ phase space of one particle. The justification for this has become clear only comparatively recently through the work of Bogoliubov, from the observation that in a gas there exist two rather widely different physical time scales corresponding to the short range of intermolecular interactions and the long mean-free path at ordinary gas densities. It is not the purpose of the present article to analyze this aspect of Boltzmann's theory or Bogoliubov's theory. Our interest is in the time-reversal property.

The Boltzmann equation is so familiar that we shall simply write it down (Boltzmann, 1876). The distribution function f is normalized according to

$$\int f(r,v,t)\,dv = n(r,t) = \text{number of molecules per unit volume}$$

The equation for $f(r,v,t)$ is

$$\frac{\partial f}{\partial t} + v \cdot \frac{\partial f}{\partial r} + \frac{K}{m} \cdot \frac{\partial f}{\partial v} = \left(\frac{\partial f}{\partial t}\right)_{\text{collision}}$$

$$= \int dv_1 \int d\omega\, \sigma |v - v_1| (f'f_1' - ff_1) \qquad (3.1.1)$$

where K is the external force, the subscript 1 denotes another molecule with which the one under consideration in $f(r,v,t)$ collides, $d\omega = \sin\theta\,d\theta\,d\phi$, θ the scattering angle, σ the collision cross section, and

$$f' = f(r,v',t), \qquad f_1' = f_1'(r,v_1',t), \text{ etc.}$$

the prime denoting quantities after a two-body collision. The first term in the integral gives the rate of increase of $f(r,v,t)$ due to the restitution collisions, and the second term gives the rate of decrease due to direct collisions. The specific form of the right-hand side constitutes the basic assumption (*Ansatz*, in German, is the more appropriate expression) of Boltzmann's theory. This *Ansatz* is not a consequence of the dynamical laws but is based on plausibility considerations. As a consequence of this *Ansatz*, the Boltzmann equation has the following properties:

(i) Equation (3.1.1) is not invariant upon the reversal of the direction of time. On reversing t as in equations (2.2.3), and writing

$$\bar{f} = f(r,-v,-t) = f(r,-v,\tau)$$

etc., one obtains the equation for f in τ which differs from that for f in t by the presence of a minus sign between the two sides of the equation. In fact this equation for f in τ does not have plausible physical meaning. Equation (3.1.1), so to speak, has a definite time arrow built in.

(ii) This definite direction of time built in in (3.1.1) is the one along which the H function

$$H(t) = \int\int f \ln f \, dr \, dv \tag{3.1.2}$$

decreases monotonically with time, for with (3.1.1), one readily finds

$$\frac{dH}{dt} = \int\int (1 + \ln f)\frac{\partial f}{\partial t} \, dr \, dv \leqslant 0 \tag{3.1.3}$$

Thus the Boltzmann equation (3.1.1) satisfies the H theorem.

Now on a comparison and identification between the results from classical thermodynamics and those from the Boltzman statistics of Section 2.3, Boltzmann obtained the following relation between the probability W of Section 2.3 and entropy

$$S = k \ln W$$

k being the Boltzmann constant.

It is also easy to obtain the following relation between the H function defined in (3.1.2) and the probability W,

$$NH = -\ln W$$

(H here being for one molecule).

Thus, up to an additive constant,

$$S = -NkH$$

The H theorem (3.1.3) thus is equivalent to the Second Law of Thermodynamics, i.e., the law of increasing entropy for the irreversible approach of a gas to equilibrium.

We must repeat that the H theorem (3.1.3) has been established by the explicit use of the Boltzman equation (3.1.1) which is *not* a dynamical law but which has been constructed to be irreversible by the *Ansatz* for $(\partial f/\partial t)_{\text{collision}}$ in (3.1.1). In this case, the criticisms of Loschmidt and of Zermelo do not apply.

3.2. *Fokker–Planck Equation* (Fokker, 1914; Planck, 1917; Chandrasekhar, 1943)

Another theory with time irreversibility is that of Fokker (1914) and Planck (1917). It is based on a less specific assumption than the collision-number *Ansatz* of Boltzmann in (3.1.1). A molecule is considered to have random collisions with other molecules. Let $Q(v - \Delta v, \Delta v)$ be the transition probability that in a time interval Δt (>0), a molecule changes its velocity from $v - \Delta v$ to v, and

$$\int Q(v - \Delta v, \Delta v) \, d\Delta v = 1 \tag{3.2.1}$$

$Q(v, \Delta v)$ depends on the present state, i.e., the velocity v, of the particle, and not on the past history. It is not an explicit function of time, but depends on the length of the interval Δt, and Δt must not be too short in order that the molecule may have suffered a sufficiently large number of random collisions with other molecules. The postulate is now made that the distribution function at time $t + \Delta t$ is related to that at time t through the Smoluchowski law

$$f(v, t + \Delta t) = \int d\Delta v f(v - \Delta v, t) \, Q(v - \Delta v, \Delta v) \qquad (3.2.2)$$

It is important to note that this is not a consequence of dynamical laws, but is an assumption connecting the probabilities at two different times irrespective of the past history before the time t.

On calculating the rate of change of $f(v)$ and introducing the averages

$$\langle \Delta v \rangle = \int \Delta v \, Q(v, \Delta v) \, d\Delta v$$

$$\langle \Delta v_x \Delta v_y \rangle = \int \Delta v_x \Delta v_y \, Q(v, \Delta v) \, d\Delta v, \text{ etc.}$$

one obtains, on writing $\partial/\partial v = \nabla_v$,

$$\left(\frac{\partial f}{\partial t} \right)_{\text{collision}} = -\nabla_v \cdot \left(\frac{\langle \Delta v \rangle}{\Delta t} f \right) + \tfrac{1}{2} \nabla_v \nabla_v : \left(\frac{\langle \Delta v \, \Delta v \rangle}{\Delta t} f \right) + \cdots \quad (3.2.3)$$

On equating this to the left-hand side of (3.1.1), one obtains the Fokker–Planck equation. One immediate application of this equation is to the Brownian motion, leading to the result obtained earlier by Einstein by other methods. We shall not go into further discussions of this equation, except to emphasize that it is not invariant upon time reversal, and that this non-invariance comes from the assumed relation (3.2.2). That this relation introduces a time arrow will be more clearly brought out in the following section.

3.3. *The Master Equation* (Wu, 1966)

Here we shall formulate a time-reversal non-invariant theory in a formal manner, leaving out detailed physical interpretation, in order to bring out the far-reaching consequences of a relation of the type (3.2.2).

Let w_i, $i = 1, 2, \ldots$, be the probability that a system be in state i at time t. Let the probabilities at a later time $t + \Delta t$ be w_k', $k = 1, 2, \ldots$. We have

$$\sum_i w_i = 1, \qquad \sum_i w_i' = 1 \qquad (3.3.1)$$

Assume that the system in state k at time t has the transition probability A_{ik}, in the interval Δt, of going into state i at time $t + \Delta t$. The transition probabilities depend on the length of the interval Δt, but are independent of the past history of the system. They are to satisfy the requirements, by virtue of their being probabilities,

$$0 \leqslant A_{ik} \leqslant 1 \qquad (3.3.2)$$

and

$$\sum_i A_{ik} = 1 \tag{3.3.3}$$

Let us make the hypothesis that relates the probabilities $w_i{}'(t + \Delta t)$ and $w_k(t)$:

$$w_i{}'(t + \Delta t) = \sum_k A_{ik} w_k(t) \tag{3.3.4}$$

It is to be noted that, except for the use of discrete states here, this relation is of the same nature as (3.2.2).

From (3.3.4), together with (3.3.3), one obtains

$$w_i{}'(t + \Delta t) - w_i(t) = \sum_k A_{ik} w_k - \left(\sum_k A_{ki}\right) w_i$$

or

$$\frac{\Delta w_i(t)}{\Delta t} = \sum_k \left(\frac{A_{ik}}{\Delta t} w_k - \frac{A_{ki}}{\Delta t} w_i\right) \tag{3.3.5}$$

which is the so-called Master Equation. If we introduce the transition probabilities per unit time

$$a_{ik} = A_{ik}/\Delta t$$

we may write (3.3.5) in the differential form

$$\frac{dw_i}{dt} = \sum_k (a_{ik} w_k - a_{ki} w_i) \tag{3.3.6}$$

The theory above has the following properties:

(i) Once the assumption (3.3.4) is made expressing $w_i{}'$ at $t + \Delta t$ in terms of the w_k at time t, it is not possible to go backward in time to find the $w_k(t)$ in terms of the $w_i{}'(t + \Delta t)$ and still to have the same probability meaning given in (3.3.4). Let us assume that the inverse matrix A^{-1} exists so that equation (3.3.4) can be solved

$$w_k(t) = \sum_i A_{ki}^{-1} w_i{}'(t + \Delta t) \tag{3.3.7}$$

From $A^{-1} A = $ unit matrix, we have

$$\sum_i A_{ki}^{-1} A_{ij} = \delta_{kj} \tag{3.3.8}$$

Take $k \neq j$. Since all A_{ij} are $\geqslant 0$, it follows from (3.3.8) that not all A_{ki}^{-1} are greater than zero. Again, take $k = j$. Since all $A_{ik} \leqslant 1$, it follows from (3.3.8) and (3.3.3) that the A_{ki}^{-1} cannot all be < 1.

From this it also follows that the Master Equation (3.3.6) has no probability meaning if one simply reverses the sign of t.

(ii) Let us define, similarly to (2.4.2a), an H function by

$$H(t) \equiv \sum_i w_i(t) \ln w_i(t) \tag{3.3.9}$$

Using (3.3.5), we have

$$\frac{dH}{dt} = \frac{1}{\Delta t} \sum_{i,k} (A_{ik} w_k - A_{ki} w_i)(1 + \ln w_i)$$

$$= \frac{1}{\Delta t} \sum_{i,k} (A_{ik} w_k - A_{ki} w_i) \ln w_i \qquad \text{on using (3.3.3)}$$

$$= \frac{1}{\Delta t} \sum_{i} (w_i' - w_i) \ln w_i \qquad \text{on using (3.3.3) and (3.3.4)}$$

By use of the same method as that used in proving the mathematical inequality (2.2.12) (which differs from the present problem in involving integration instead of summation), it can be seen that the right-hand side of the last line in the equations above is negative. Hence

$$\frac{dH}{dt} \leqslant 0 \qquad\qquad\qquad (3.3.10)$$

This shows that the theory has a definite direction of time.†

† The proof of the relation (3.3.10) can be carried out without appealing to (2.2.12) but by using an extra normalization for the A_{ik}, namely,

$$\sum_{k} A_{ik} = 1 \qquad\qquad\qquad (3.3.3a)$$

The meaning of this condition is not as obvious as (3.3.3). But both (3.3.3) and (3.3.3a) are suggested in a quantum mechanical version of the theory where the relations (3.3.3) and (3.3.3a) are properties of the unitary transformation in time (Wu, 1966).
   We calculate, from (3.3.4) and (3.3.9),

$$H(t + \Delta t) - H(t) = \sum_{i,k} A_{ik} w_k \ln w_i' - \sum_{i} w_i \ln w_i$$

Now we have the following inequality due to Gibbs. For any positive x, we have

$$\int_{1}^{x} \ln y \, dy = x \ln x - x + 1 \geqslant 0$$

Set $x = w_k/w_i'$, multiply the above inequality relation by w_i' and we get

$$w_k \ln w_k - w_k + w_i' \geqslant w_k \ln w_i'$$

Since all $A_{ik} \geqslant 0$, we obtain

$$H(t + \Delta t) - H(t) \leqslant \sum_{i,k} A_{ik}(w_k \ln w_k - w_k + w_i') - \sum_{i} w_i \ln w_i$$

On using (3.3.3) and (3.3.3a), it is seen that the right-hand side is zero, and we obtain

$$H(t + \Delta t) - H(t) \leqslant 0$$

   It might have appeared at first thought that had one used the inverse relation (3.3.7), all the above steps could have been repeated, with A_{ik}^{-1} replacing A_{ik}, and one would have arrived at

$$H(t) - H(t + \Delta t) \leqslant 0$$

and this would have meant

$$H(t + \Delta t) - H(t) = 0$$

This result, however, it not true, because the A_{ik}^{-1} are *not* all >0 as we have shown before, and one cannot arrive at the inequality $H(t) - H(t + \Delta t) \leqslant 0$.

(iii) The Master Equation, subject to the basic assumption (3.3.4), is quite general. By replacing the discrete states by continuous states, the H function defined in (3.3.9) can be compared with the H function in (2.2.5). It may be recalled that on the basis of the dynamical laws, the H function in (2.2.5) remains unchanged in time, whereas on the basis of the probability assumption (3.3.4), the H function in (3.3.9) decreases monotonically with time.

(iv) The linearity of the Master Equation is only apparent, because the transition probabilities a_{ik} can depend on the probabilities w_k. In fact the Boltzmann equation (3.1.1) can be thrown into the form (3.3.6) in which the transition probabilities a_{ik} are proportional to the $w_i = f(r, vi, t)$ in (3.1.1). One may regard the Boltzmann equation as a special case of the Master Equation. They are both time-reversal non-invariant.

3.4. Theories Based on the Liouville Equation

As seen in Section 2, the Liouville equation is time-reversal invariant and it is not possible to obtain an irreversible equation from it without additional assumptions. But it is also seen that the theories of Boltzmann and Gibbs make the approach of a gas to equilibrium extremely probably irreversible. In recent years many attempts have been made to formulate a theory of gases (and plasmas) by starting from the Liouville equation. The common initial step in many of these theories is to transform the Liouville equation into a system of coupled equations by defining s-particle distribution functions

$$F_s(x_1, x_2, \ldots, x_s, t)$$

(in which x_i stands for q_i, p_i, $s = 1, 2, \ldots, N$) by

$$F_s = V^s \int \cdots \int F_N \, dx_{s+1}, \ldots, dx_N, \qquad F_N \text{ proportional to } \rho \text{ in (2.2.1)}$$

V is the volume of the gas. Let ϕ_{ij} denote the intermolecular interaction between particles i and j. The Hamiltonian of an s-particle subsystem of the gas is

$$H_s = \sum_i \left(\frac{1}{2m} p_i^2 + U(q_i) \right) + \sum_{1 \leqslant i < j} \phi_{ij}$$

Let us introduce the Poison bracket expression

$$\{A, B\} = \sum_i \left(\frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial B}{\partial q_i} \right)$$

Then simply by integration of the Liouville equation (2.2.2), one obtains the system of coupled differential-integral equations

$$\frac{\partial F_s}{\partial t} - \{H_s, F_s\} = \frac{N - s}{V} \int dx_{s+1} \left\{ \sum_{i=1}^{s} \phi_{i, s+1}, F_{s+1} \right\}, \qquad s = 1, 2, \ldots \quad (3.4.1)$$

which is known as the B–B–G–K–Y hierarchy (Bogoliubov, 1946; Born & Green, 1946, 1947; Kirkwood, 1946, 1947; Yvon, 1935). This system of equations is completely equivalent to the Liouville equation and is time-reversal invariant.

Different theories differ at this point in their different expansions of the F_s and the different methods of introducing a time-irreversibility into the theory. It is beyond the scope of the present article to describe the various theories. We wish only to emphasize the time-irreversibility aspect of these theories.

In the theory of Bogoliubov, a time direction is introduced by the 'initial condition' which *defines* that direction as the past in which the s-particle distribution function F_s is uncorrelated, i.e.,

$$\lim_{t \to -\infty} F_s = F_1(1) F_1(2), \ldots, F_1(s), \qquad s = 2, 3, \ldots, N \qquad (3.4.2)$$

The positive direction of time, or the future, is the one along which the particles become correlated on account of their interactions. When the conditions (3.4.2) is used in integrating (3.4.1), the result is no longer time-reversal invariant. It might be mentioned in passing that by expanding the F_s in powers of the gas density and terminating the equation for F_1 from (3.4.1) at the first order in density, Bogoliubov obtained an equation which, in some approximations, reduces to Boltzmann's equation (3.1.1). This is a most satisfactory feature of Bogoliubov's theory, since it furnishes a foundation for the Boltzman equation. A similar theory of obtaining Boltzmann's equation from the Liouville equation has been independently given by Kirkwood.

In the theories of other authors, either an initial condition on the many-particle function F_s is used, or a Laplace transform in time is made

$$f(k, v, \omega) = \int_0^\infty f(k, v, t) \exp(-\omega t) \, dt \qquad (3.4.3)$$

where the real part of ω is positive. Such a transform automatically excludes negative values of t from the theory, and only positive values of t are permissible.

4. Concluding Remarks

From the preceding sections, we can summarize the discussions as follows.

(1) There are the observed irreversible processes in gases. If we are interested only in formulating laws to describe them in the macroscopic view, then we have already satisfactory laws.

(2) But if we are interested in formulating theories from the atomic level, i.e., in understanding the macroscopic irreversibility on the basis of

22

dynamical laws governing the motions of the molecules, then there are two major problems. One is the impracticability, and irrelevance, of dealing with the large number of molecules in any macroscopic amount of gas. Probability concepts and statistical methods are introduced to deal with this aspect. The other problem is to get out, from the fundamentally reversible dynamical laws, theories that will describe the observed irreversible processes.

(3) The marriage between probability concepts and dynamics leads to the Liouville equation which however is dominated by dynamics and is time-reversible invariant. Without the addition of other assumptions outside of dynamical laws, one cannot obtain from the Liouville equation a theory which is time-irreversible explicitly. In particular, one cannot establish such an irreversibility as the H theorem on purely dynamical ground.

(4) On the other hand, on the basis of the Liouville equation, it is possible to understand the extremely high probability of 'irreversible' approach of a gas toward equilibrium. This is achieved in the theory of Boltzmann for one single system and that of Gibbs for an ensemble. Thus the Liouville equation, although mathematically reversible in the time, does contain the essence of 'physical' or 'macroscopic' irreversibility in the probability sense. This property of the Liouville equation has its origin in the largeness of the number of particles in the system under consideration.

(5) Because of this last-named property of the Liouville equation, it follows that any theory based on it and rendered time-reversal non-invariant by any plausible assumption or artifice will also describe the irreversible approach to equilibrium. This is illustrated by the theories of Bogoliubov and others who introduce certain initial conditions that in essence define a time arrow. In such theories, the objections of Loschmidt and of Zermelo will not arise.

(6) Now we may ask ourselves the question as to what we really want regarding a theory for irreversible processes. If we allow ourselves to be less puritanical as far as strict adherence to dynamical laws is concerned, then our object is to formulate a theory with the purpose of describing the irreversible processes. We may retain dynamics as far as possible but are ready to introduce additional ideas which are not deducible from dynamics.

It is on this philosophy that many theories have been proposed. Boltzmann's equation, the Fokker–Planck equation and the Master Equation are briefly referred to in Sections 2 and 3.

(7) There are authors who are not happy with this attitude. Some try to trace this irreversibility (in the macroscopic scale) to something more basic. It seems that at the moment not much can be reported about such ideas.

References

Bogoliubov, N. N. (1946). *Fiziologicheskiĭ Zhurnal SSSR im. I. M. Sechenova*, **10**, 265. Monograph translated in *Studies in Statistical Mechanics*. Edited by de Boer and Uhlenbeck (1962), Interscience.

Boltzmann, L. (1872). *Sitzungsberichte der Akademie der Wissenschaften in Wien*, **66**, 275.

Boltzmann, L. (1875). *Sitzungsberichte der Akademie der Wissenschaften in Wien*, **72**, 427.

Boltzmann, L. (1876). *Sitzungsberichte der Akademie der Wissenschaften in Wien*, **74**, 503.

Born, M. and Green, H. S. (1946). *Proceedings of the Royal Society (London)*, **A188**, 10.

Born, M. and Green, H. S. (1947). *Proceedings of the Royal Society (London)*, **A189**, 103; **A190**, 455; **A191**, 168.

Chandrasekhar, S. (1943). *Reviews of Modern Physics*, **15**, 1.

Fokker, A. D. (1914). *Annalen der Physik*, **43**, 812.

ter Haar, D. (1954). *Elements of Statistical Mechanics*. Rhinehart & Co.

ter Haar, D. (1955). *Reviews of Modern Physics*, **27**, 289.

Kirkwood, J. G. (1946). *Journal of Chemical Physics*, **14**, 180.

Kirkwood, J. G. (1947). *Journal of Chemical Physics*, **15**, 72.

Liouville, J. (1838). *Journal de mathématiques pures et appliquées*, **3**, 348. See also, Whittaker, E. T. (1927). *Analytical Dynamics*. Cambridge University Press.

Loschmidt, J. (1876). *Sitzungsberichte der Akademie der Wissenschaften in Wien*, **73**, 139.

Loschmidt, J. (1877). *Sitzungsberichte der Akademie der Wissenschaften in Wien*, **75**, 67.

Onsager, L. (1931). *Physical Review*, **37**, 405.

Onsager, L. (1931). *Physical Review*, **38**, 2865. See also Casimir, H. B. G. (1945). *Reviews of Modern Physics*, **17**, 343, for a proof of the reciprocity relations on the principle of microscopic reversibility.

Planck, M. (1917). *Sitzungsberichte der Preussischen Akademie der Wissenschaften zu Berlin*, **5**, 324.

Poincaré, H. (1890). *Acta mathematica, Stockholm*, **13**, 67.

Tolman, R. C. (1935). *Principles of Statistical Mechanics*. Oxford University Press.

Wu, T. Y. (1966). *Kinetic Equations of Gases and Plasmas*. Addison-Wesley.

Yvon, J. (1935). *La Theore Statistique des Fluids et d'Equation d'Etat*. Paris.

Zermelo, E. (1896). *Annalen der Physik*, **57**, 485.

For general references, in addition to the works of Tolman, R.C. (1935) and ter Haar, D. (1954), see Uhlenbeck, G. E. and Ford, G. W. (1963). *Lectures in Statistical Mechanics*. American Mathematical Society, Providence, Rhode Island; and the classic work of Ehrenfest, P. and Ehrenfest, T. (1911). *Enzyklopädie der mathematischen Wissenschaften*, Vol. IV, Part 32 (English translation published by Cornell University Press, 1959).