International Journal of Theoretical Physics, Vol. 14, No. 5 (1975), pp. 289–294

Boltzmann's *H* Theorem and the Loschmidt and the Zermelo Paradoxes

TA-YOU WU

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

Received: 24 *February* 1975

Abstract

The Umkehreinwand of Loschmidt and the Wiederkehreinwand of Zermelo have been reexamined. The former paradox depends on the augument that for a dynamical system, upon the reversal of the velocities of all the molecules, the H function retraces its sequence of values so that dH/dt will change its sign. The latter paradox depends on the argument that the H function returns infinitely close to its value after a Poincare' quasiperiod and therefore cannot be decreasing all the time. While the main contention of the two paradoxes is correct, that the H theorem is inconsistent with classical dynamical laws, the arguments there can be considerably simplified and the "paradoxes" answered more directly. If the distribution function $f(q_K, p_K, t)$ is governed by an equation which is time-reversal invariant (such as the Liouville equation for a closed dynamical system), then it can be shown immediately that $dH/dt = 0, H = \text{cons}$. In this case, both paradoxes disappear, but together with them, the $dH/dt < 0$ part of the H theorem also has disappeared, i.e., there is no second law of thermodynamics. If $f(q_K, p_K, t)$ is governed by an equation which is not time-reversal invariant (such as the Boltzmann equation, or the Master Equation for Markovian processes), then (1) there is no argument for f and $H(t)$ to retrace their sequence of values upon the reversal of all the velocities of the system, (2) there is no quasiperiod in which f and $H(t)$ return to their earlier values. In this case, both paradoxes disappear also, but then one must go beyond classical dynamics in order to maintain the H theorem.

1. *Introduction*

The H theorem states that the H function defined by

$$H(t) = \int f(q_K, p_K, t) \ln f(q_K, p_K, t)\, dq_K dp_K \tag{1.1}$$

where $q_K, p_K = q_1, \ldots, q_n, p_1, \ldots, p_K$, satisfies the relation

$$dH(t)/dt \leqslant 0 \tag{1.2}$$

The paradoxes of Loschmidt (1876) and of Zermelo (1896) are the most famous of many discussions of the theorem. The former argues as follows:

Consider a system A whose phase (q_K, p_K) passes through the points $P_1, P_2, P_3, \ldots$ of the Γ space at time $t_1, t_2, t_3, \ldots$. According to (1.2), we have

$$H_A(t_1) \geqslant H_A(t_2) \geqslant H_A(t_3) \ldots, \quad t_1 < t_2 < t_3 \cdots \tag{1.3}$$

Consider another system B which differs from A only in having all the velocities opposite to those of A, i.e.,

$$q'_K = q_K, \qquad p'_K = -p_K$$

The phase (q'_K, p'_K) of B passes through the points $\cdots P'_3, P'_2, P'_1$ at time $t'_3 < t'_2 < t'_1 \cdots P'_n$ and P_n are correspond points in the sense that

$$P_n = (q_K, p_K), \qquad P'_n = (q_K, -p_K) \tag{1.4}$$

and according to (1.1),

$$H_A(t_n) = H_B(t'_n) \tag{1.5}$$

and from (1.3)

$$H_B(t'_1) \geqslant H_B(t'_2) \geqslant H_B(t'_3) \ldots, \quad t'_3 < t'_2 < t'_1 \cdots \tag{1.6}$$

which means

$$dH_B/dt \geqslant 0 \tag{1.7}$$

This is contradictory to (1.2).

The Zermelo (1896) argument is as follows: According to Poincaré's ergodic theorem, the phase (q_K, p_K) comes back infinitely near any given initial point in the Γ space after a sufficiently long time, the quasiperiod T, and the trajectory passes through points $P'_1, P'_2, P'_3, \ldots$ infinitely close to $P_1, P_2, P_3, \ldots$ at $t_1 + T, t_2 + T, t_3 + T, \ldots$. Zermelo argues that

$$H(t_1 + T) \cong H(t_1), \qquad H(t_2 + T) \cong H(t_2), \qquad \text{etc.} \tag{1.8}$$

so that there is the paradox that H has kept on decreasing according to (1.3) and yet after a quasiperiod gets back to an earlier value (1.8).

These and other criticisms of the H theorem in the original form have led Boltzmann to reinterpret the theorem on a probability basis. Thus the Stosszahlansatz in the Boltzmann equation is interpreted as giving an overwhelmingly large probability for H to decrease, and fluctuations in which H increases are allowed but are not explicitly contained in the Boltzmann equation.

The many discussions on the H theorem have been summarized by the Ehrenfests (1911), Tolman (1935), and ter Haar (1954, 1955). The Ehrenfests in particular discuss the fluctuations of H in discrete steps due to molecular collisions.

2. *Comments on the Loschmidt and Zermelo Paradoxes*

On looking back at these two famous paradoxes, it seems that these criticisms can be answered in a more basic way. Consider first the nature of the equation that governs the variation of $f(q_K, p_K, t)$ with time. If the equation

$$df/dt = I(f) \tag{2.1}$$

is invariant with respect to time reversal, that is the transformation

$$t \to -\tau,\ f(q_K, p_K, t) \to f(q_K, -p_K, -t- = \bar{f}(q_K, p_K, \tau) \tag{2.2}$$

transforms (2.1) into

$$d\bar{f}/d\tau = I(\bar{f}) \tag{2.3}$$

then it can be shown that $I(f) = 0$ and therefore[1]

$$dH/dt = 0, \quad \text{or } H = \text{cons.} \tag{2.4}$$

[1] The proof of (12) is as follows. Upon velocity reversal, $p \to -p$, let us denote

$$f(q, p, t) \to f(q, -p, t) \equiv f'(q, p, t) \tag{a}$$

Equation (a) must hold for arbitrary velocities, i.e.,

$$\frac{df}{dt} = I(f) \to \frac{\partial f'}{\partial t} = I(f') \tag{b}$$

or

$$I(f) \to I(f') \tag{c}$$

Upon t reversal, we have (2.2), which can be expressed as

$$f(q, p, t) \to f(q, -p, -t) = f'(q, p, -\tau) \equiv \bar{f}(q, p, \tau) \tag{d}$$

By hypothesis, (2.1) is invariant with respect to t reversal,

$$\frac{df}{dt} = I(f) \to \frac{d\bar{f}}{d\tau} = I(\bar{f}) \tag{e}$$

i.e.,

$$I(f) \to -I(\bar{f}) \tag{f}$$

Let the t reversal be made at $t = \tau = 0$. Then (d) gives

$$\bar{f}(q, p, o) = f'(q, p, o) \tag{g}$$

so that at $t = \tau = 0$, (f) gives

$$I(f(q, p, o)) \to -I(f\ (q, p, o)) \tag{h}$$

Comparison with (c) gives

$$I(f) = 0 \tag{i}$$

For a closed system governed by classical dynamical laws, Eq. (2.1) with $I(f) = 0$ is the Liouville equation for which (2.4) can be obtained immediately.

If $f(q_K, p_K, t)$ is governed by an equation which is invariant under the velocity reversal operation, i.e.,

$$p_K \to -p_K, \quad f(q_K, p_K, t) \to f(q_K, -p_K, t) \equiv \bar{\bar{f}}(q_K, p_K, t) \tag{2.5}$$
$$df/dt = I(f) \to d\bar{\bar{f}}/dt = I(\bar{\bar{f}}), \quad H(|f) \to H(|\bar{\bar{f}}) = H(|f)$$

then it can immediately be seen that

$$\frac{dH(|\bar{\bar{f}})}{dt} = \frac{dH(|f)}{dt} \tag{2.6}$$

Invariance with respect to velocity reversal must be assumed in any plausible theory which is to apply to *all* possible velocities for the molecules. The Boltzmann equation, for example, satisfies this invariance requirement. The time-reversal invariance is, however, a much stronger condition and is not satisfied by the Boltzmann equation.

With these general considerations, we are in a position to examine the two paradoxes. Consider the Loschmidt criticism.

(i) If $f(q_K, p_K, t)$ is governed by a time-reversal invariant equation (2.1) and (2.3), then we simply have (2.4)

$$dH/dt = 0$$

and H is a constant, so that the inequality (1.7) disappears, together with the $dH/dt < 0$ part of the H theorem $dH/dt \leqslant 0$, i.e., there is no law of entropy increase.

(ii) If $f(q_K, p_K, t)$ is governed by a velocity-reversal invariant equation (2.5), then we must have (2.6), i.e., if

$$dH(|f)/dt \leqslant 0$$

the velocity-reversed system must also have

$$dH(|\bar{\bar{f}})/dt \leqslant 0 \tag{2.7}$$

(iii) The basis (1.5) of the paradox is fallacious. If $dH/dt \neq 0$, the equation governing $f(q_K, p_K, t)$ cannot be time-reversal invariant, and reversing the velocities will not cause f to retrace its sequence of values even when the phase (q'_K, p'_K) passes through $P'_3, P'_2, P'_1 \cdots$ at $t'_3 < t'_2 < t'_1 \cdots$ where P_n and P'_n correspond as in (1.4). $f(q_K, p_K, t)$ changes not only in value, but also in the functional form, with time and in an irreversible manner [since the equation $\partial f/\partial t = I(f)$ is not reversible]. Without (1.5), there is no paradox.

Consider next the Zermelo paradox. Here again: Either (i) $f(q_K, p_K, t)$ is governed by a time-reversal invariant equation, in which case we have (2.4) and H is a constant, so for this case (1.8) is valid, but the paradox of H decreasing and getting back to an earlier value after a quasiperiod disappears, or (ii) $f(q_K, p_K, t)$ is governed by an irreversible or stochastic equation in which case f changes in time both in value and in functional form,[2] and in general there is no quasiperiod after which $f(q_K, p_K, t)$ and $H(t)$ return to their earlier value. In this case, the relations (1.8) are not true and there is again no paradox. But then $dH/dt \leqslant 0$ is not a consequence of dynamical laws.

3. *Concluding Remarks*

To sum up, we believe that the many involved discussions of the H theorem can be simplified with our present understandings.

(i) On classical dynamical laws, a closed dynamical system can have only

$$H = \text{const}$$

i.e., there is no law of increasing entropy. This is the contention of Loschmidt and Zermelo. The arguments in the two paradoxes are greatly simplified by the general result (2.4).

(ii) The $f(q_K, p_K, t)$ in the H function (1.1) satisfying the H theorem (1.2) cannot be governed by an equation which is a consequence of reversible dynamical laws; it could be an equation based on general probability assumptions (for example, the Smoluchowski assumption for Markovian processes leading to the Master equation, or the Fokker-Planck equation) or on specific probability assumptions (such as the Boltzmann equation), or on other assumptions. In such cases, the contentions (1.5) and (1.8) in the paradoxes are not valid.

(iii) In the face of these paradoxes and other criticisms, Boltzmann later interpreted the H theorem (1.2) on probability basis, that is, whenever H is above its minimum value, it has an overwhelmingly large probability to decrease, but an increase of H due to molecular collisions is not absolutely ruled out. In the particular case of the Boltzmann equation, the Stosszahlansatz gives this large probability and (1.5) and (1.8) are not relevant.

(iv) The really significant refinement of the original form of the H theorem is Boltzmann's later probability interpretation and the implied extension to include fluctuations. Recently Fox & Uhlenbeck (1969) have suggested the addition of a fluctuation term to the Boltzmann equation. Lee & Wu (1973), starting with the Liouville equation in the form of the *B–B–G–K–Y* hierarchy, have obtained such a fluctuation term on statistical consideration of the many-particle correlations. From a Boltzmann equation with fluctuation, one obtains immediately an H function which decreases with time in the main but has fluctuations.

References

Ehrenfest, P., and Ehrenfest, T. (1911). Article in *Enzyklopädie der Wessenschaften* IV, (32). (English translation, Cornell University Press, 1959).

[2] For example, the velocity distribution of a gas may start with any arbitrary initial function and approach the Maxwell distribution in time.

Fox, R. F., and Uhlenbeck, G. E. (1969). "Contributions to the Theory of Non-equilibrium Thermodynamics" (Dissertation), Rockefeller University, New York.
Lee, Paul S., and Wu, T. Y. (1973). *International Journal of Theoretical Physics,* 7, 267.
Loschmidt, L. (1867). *Sitznngsberichte der Akademie der Wissenschaften in Wien*, **73**, 139; **75**, 67.
ter Haar, D. (1954). *Elements of Statistical Mechanics*, Appendix I. Rinehart, New York.
ter Haar, D. (1955). *Reviews of Modern Physics*, **27**, 289.
Tolman, R. C. (1935). *Principles of Statistical Mechanics*, Oxford University Press, New York.
Zermelo, E. (1896). *Annalen der Physik*, **57**, 485; **59**, 793.