

CHAPTER 6

BOLTZMANN'S H-THEOREM

In the latter part of the nineteenth century, Ludwig Boltzmann almost single-handedly established the field now known as statistical mechanics. One of his major contributions, and undoubtedly the most controversial, was his H-theorem. This work was published in 1872 with the intent of showing that the second law of thermodynamics derives from the laws of mechanics. While it is now generally agreed that the attempted proof was unsuccessful, the controversy attending this work was ultimately beneficial to the burgeoning field of statistical mechanics because it forced the workers to think through and refine the statistical and probabilistic concepts that were introduced by Boltzmann.

6.1 THE H-THEOREM

$$H = \sum_i f_i \ln f_i \delta x_i \delta y_i \delta z_i \delta p_{x_i} \delta p_{y_i} \delta p_{z_i} \quad (6-1)$$

Boltzmann began by defining the function H for a dilute gas comprised of spherical particles¹ where f is a distribution function which determines the number of particles n_i located in the spatial region $\delta x_i \delta y_i \delta z_i$ and having momentum in the range $\delta p_{x_i} \delta p_{y_i} \delta p_{z_i}$ through the relation

$$n_i = f(x, y, z, p_x, p_y, p_z, t) \delta x_i \delta y_i \delta z_i \delta p_{x_i} \delta p_{y_i} \delta p_{z_i} \quad (6-2)$$

The term $\delta x_i \delta y_i \delta z_i \delta p_{x_i} \delta p_{y_i} \delta p_{z_i}$ is denoted $\delta v_{\mu i}$ and is referred to as the "volume" of a cell in 6-dimensional μ -space. The cells occupy equal "volumes" of μ -space. Each particle has six degrees of freedom and could be completely specified by a point in μ -space. Thus, a quantity of gas containing N particles can be represented by a swarm of N points in μ -space and the distribution function f tells us how these N points are partitioned among the cells of μ -space. The summation in Eq. (6-1) is taken over all of the cells in μ -space. As indicated in

¹ This derivation closely follows R. C. Tolman, *The Principles of Statistical Mechanics*, Oxford University Press, London, 1938, pp. 134-140.

Eq. (6-2), the distribution function could depend upon position, momentum, and time.

The function H can be restated as²

$$H = N \sum_i \frac{n_i}{N} \ln \frac{n_i}{N} + constant \quad (6-3)$$

If n_i/N could be taken as the probability of a particle being found in the i th cell of μ -space, we could write Eq. (6-3) as

$$H = N \sum_i P_i \ln P_i + constant \quad (6-4)$$

The first right-hand term of Eq. (6-4) would appear to be related to the statistical mechanical entropy³, but it must be remembered that the latter quantity refers to an equilibrium state and therefore the P_i 's should be the cell occupation probabilities when the equilibrium distribution prevails. Our substitution of n_i/N for P_i implies that the following relationship between H and S is valid only as equilibrium is approached

$$H = - \frac{S}{k} + constant \quad (6-5)$$

or

$$-k \frac{dH}{dt} = \frac{dS}{dt} \quad (6-6)$$

The time derivative of S can therefore be obtained from the time derivative of H which in turn depends on the change in f with time. Particle collisions provide the mechanism for changes in f and when molecular chaos is assumed, it can be shown that⁴

$$\frac{dH}{dt} \leq 0 \quad (6-7)$$

² See Appendix 6A for details.

³ Compare with Eq. (2-10).

⁴ For details see Appendix 6B.

As per Eq. (6-6) this results in

$$\frac{dS}{dt} \geq 0 \quad (6-8)$$

Thus, H can never increase and if Eq. (6-5) is valid, S can never decrease. These derivatives become zero at equilibrium where together forward and reverse collisions zero out and the Maxwell-Boltzmann distribution prevails.⁵

6.2 THE PARADOXES

The scientific community was initially skeptical of Boltzmann's result and a lengthy controversy ensued.⁶ For the most part, the attacks on the H-theorem centered on two features which are now known as paradoxes associated with the names of Zermelo and Loschmidt.

Zermelo's paradox, also known as the recurrence paradox, calls Boltzmann's result into question because, according to a theorem proved by Poincaré: any mechanical system of fixed energy and volume obeying the laws of classical mechanics (the type of system considered by Boltzmann) must eventually return arbitrarily close to its initial state. Thus, Boltzmann's H should eventually increase as the system returns to its initial state and therefore H can not be said to never increase. It turns out that the key word here is eventually; Poincaré's theorem specifies a finite recurrence time, but estimates show this time to be astronomically large. Thus, while accepting the validity of Zermelo's claim, Boltzmann was able to argue that because of very long recurrence times, no one would ever observe a system with increasing H and therefore his H-theorem would not conflict with experience. Boltzmann also used the argument that the Second Law is statistical in nature and stated that while systems with increasing H were possible, their occurrence could be assigned an extremely low probability.

There does not seem to be a single precise statement of the paradox attributed to Loschmidt. In his only published statement on the subject,

⁵Defined by Eq. (2-4).

⁶ A detailed account is provided by S.G. Brush, *The Kind of Motion We Call Heat*, North-Holland Publishing Co., Amsterdam, 1976, pp. 598-612.

Loschmidt wrote⁷ "...the entire course of events will be retraced if at some instant the velocities of all its parts are reversed." Boltzmann in his response associated the objection with Loschmidt and because the two were colleagues, there seems to be little doubt that they had discussed the matter. The gist of Loschmidt's objection was that Boltzmann had used Liouville's equation, which is based on the classical equations of motion, to describe the behavior of his system. Because these classical equations are symmetric in time and because the use of statistical methods introduces no asymmetry, it is not reasonable to expect the asymmetric behavior of Boltzmann's H . Thus, Loschmidt's paradox might be simplified to: How is it possible to obtain time-asymmetric behavior for a system that obeys time-symmetric equations? This question was not settled decisively until van Kampen⁸ in 1962 stated that the molecular chaos assumption used by Boltzmann in determining collision frequencies provides the source of the time-asymmetric behavior. Thus, molecular collisions are represented by a Markovian process where the history of the system is forgotten or ignored and does not determine future behavior.

It is easy to demonstrate that a simple Markovian process shows the same type of time-asymmetric behavior as Boltzmann's H . We will use the famous dog-and-fleas problem. There are two dogs A and B that share a total of N fleas serially numbered from 1 to N . Also, there is an urn containing balls numbered from 1 to N from which balls are drawn at random. When a ball is drawn, the number is read and the corresponding flea jumps from the dog it is on to the other dog. The ball is then returned to the urn, mixed with the other balls, and the drawing continues.

The situation in which all fleas are initially on dog A has been computed for the case of $N=500$. The progress of the game is shown on Fig. 6-1 where the number of fleas on each dog is plotted versus the number of events (draws). A random number generator simulated the drawing of balls. Also plotted on Fig. 6-1 is a quantity labeled "*Entroflea*" defined as

$$"Entroflea" = -(f_A \ln f_A + f_B \ln f_B)$$

⁷ *ibid.*

⁸ N.G. van Kampen in *Fundamental Problems in Statistical Mechanics*, E.G.D. Cohen ed., North-Holland Publishing Co., Amsterdam, 1962, pp. 173-182.

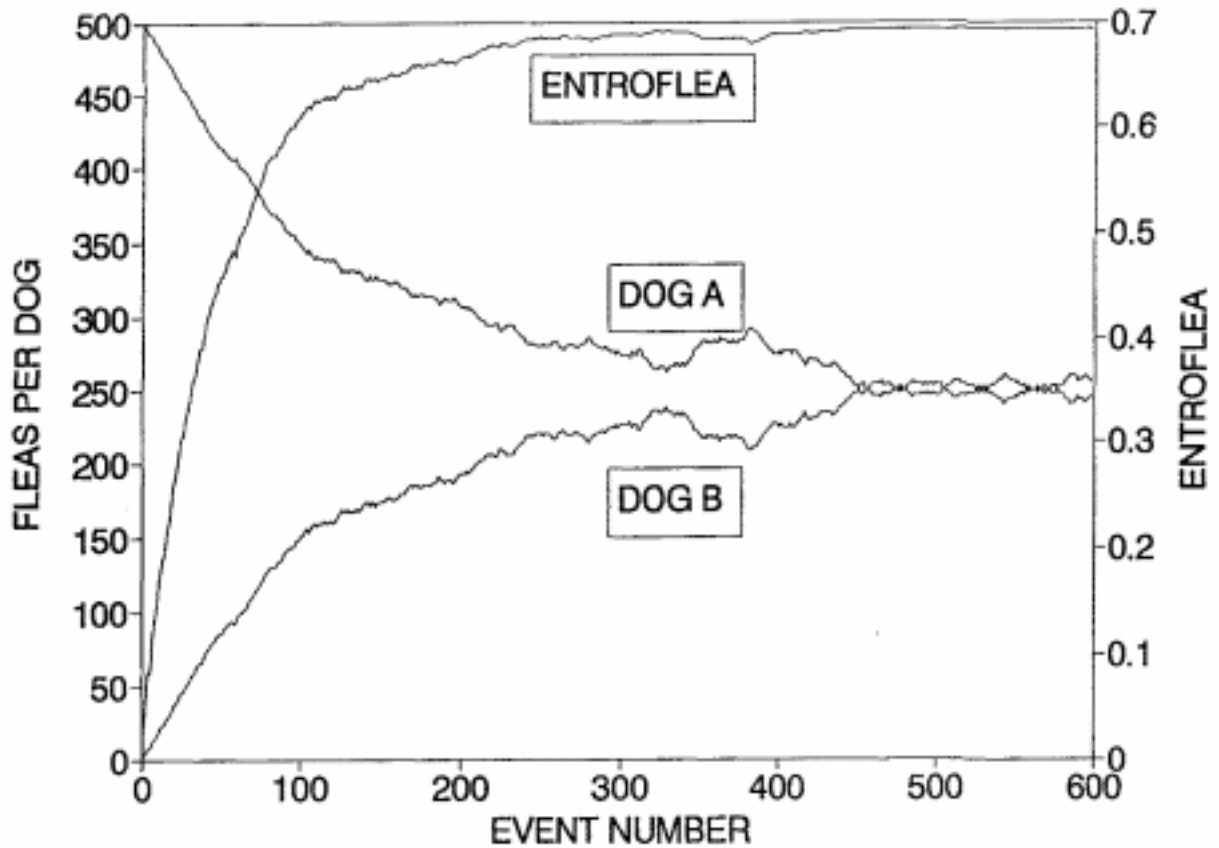


Figure 6-1. Dogs and fleas— a Markovian process.

where f_A and f_B are the fractions of the total number of fleas on dogs *A* and *B* respectively. As suggested by Eq. (6-3), "*Entroflea*" is analogous to $-H$. Note that despite minor fluctuations, "*Entroflea*" increases to an asymptotic value in much the same way as the H-theorem seems to predict for the entropy. This is characteristic of a Markovian process where each event is determined solely in terms of probability, but there is no way "*Entroflea*" can be related to the thermodynamic entropy which is defined in terms of a reversible heat effect. Thus, irreversibility is not explained by the H-theorem, but sneaks into the derivation through the necessary but deceptively innocent assumption of molecular chaos.

6.3 COMMENTARY

When considering the import of Boltzmann's H-theorem, there are two major areas of disappointment: the inability to directly relate H to the entropy except very near the final or equilibrium state, and the failure to identify the source of irreversibility.

It has already been noted that Eq. (6-5) is valid only at equilibrium, after H has ceased to change. We can therefore expect that the result $dS/dt \geq 0$ applies only near equilibrium and may wish to inquire as to the type of process for which H (or rather dH/dt) has been evaluated. In obtaining the result stated by Eq. (6-7), the distribution function f was assumed independent of spatial position and therefore we must recognize that we have been considering a macroscopically homogeneous gas. We then ask whether our calculated change in H corresponds to an observable change in state for which we expect the entropy to change. We conclude that a change of state, occurring in an isolated system in which the gas remains macroscopically homogeneous, is difficult to imagine. In this regard, ter Haar⁹ in his treatment of the H-theorem states that "if a distribution differs appreciably from the equilibrium distribution, the return to equilibrium is quite rapid." His estimate of the relaxation time for this process in a gas at 300 K and atmospheric pressure is 10^{-9} seconds which clearly indicates that the changes we have considered for H do not correspond to observable changes in entropy. Thus, the H-theorem has little to do with entropy and mainly demonstrates the stability of the equilibrium distribution.

An approach employing distribution functions and the molecular chaos assumption produces good results when applied to the calculation of transport properties and it seems reasonable that Boltzmann's H , if measurable, would indeed be found to never increase. Also, it is well known that the Sackur-Tetrode equation, a result of equilibrium statistical mechanics, correctly predicts entropy changes between well-defined changes of state of an ideal gas. Yet, in spite of these successes in the equilibrium and non-equilibrium realms, the origin of or the contributive mechanism for irreversibility has not been found. Some have proposed a cosmic rather than a local origin.¹⁰

⁹ D. ter Haar, *Elements of Thermostatistics*, Holt, Rinehart, and Winston, New York, 1966, 2nd ed., pp. 20-22.

¹⁰ See, for example, B. Gal-Or, in *A Critical Review of Thermodynamics*, E.B. Stuart, A.J. Brainard, and B. Gal-Or eds., Mono Book Corp., Baltimore, 1970, pp. 445-472.

The failure of molecular theory to provide a proof or explanation of the second law may be due to incompatible descriptions of reality.¹¹ In the molecular view, ideality is found in the motion of particles which, as we have seen, is reversible and operates in a timeless fashion independent of human presence. On the other hand, ideality in thermodynamics is represented by the reversible process which never occurs naturally and can be approached only through the intervention of a human agent in reducing friction and potential gradients. Thus, Boltzmann's H-theorem may have failed because congruence is not possible in the mapping of one view of reality onto the other.

¹¹ See Sec. 9.3

APPENDIX 6A
Simplification of H

The function H can be shown to be related to the statistical mechanical entropy. We begin by rewriting Eq. (6-2)

$$f_i = n_i / \delta v_\mu$$

and substituting this result into Eq. (6-1) to obtain

$$H = \sum_i \frac{n_i}{\delta v_\mu} \ln \frac{n_i}{\delta v_\mu} \delta v_\mu$$

which reduces to

$$H = \sum_i (n_i \ln n_i - n_i \ln \delta v_\mu)$$

or

$$H = \sum_i (n_i \ln n_i - N_i \ln \delta v_\mu)$$

Because the cell volume δv_μ is constant, the second right-hand term in this equation is constant. A bit of algebraic manipulation yields

$$H = N \sum_i \frac{n_i}{N} \ln \frac{n_i}{N} + [N \ln N - N_i \ln \delta v_\mu]$$

where the right-hand bracketed term is a constant.

APPENDIX 6B

The effect of binary collisions upon H

Before determining dH/dt , it is convenient to regard f as a continuous function in μ -space and restate Eq. (6-1)

$$H = \int \dots \int f \ln f \, dv_\mu$$

Because integration is over all six coordinates of μ -space, the integral is a function only of time. Thus, we may write

$$\frac{dH}{dt} = \int \dots \int \left[\frac{df}{dt} \ln f + \frac{df}{dt} \right] dv_\mu \quad (6-9)$$

Remembering that $f dv_\mu$ is the number of particles in a cell of differential size and noting that the integral of f over all of μ -space simply results in the total number of particles, we find the second term of the integral to be zero and write Eq. (6-9) as

$$\frac{dH}{dt} = \int \dots \int \frac{df}{dt} \ln f \, dv_\mu \quad (6-10)$$

Further, if we assume f to be independent of position, we can integrate over the spatial coordinates x , y , and z to obtain

$$\frac{dH}{dt} = V \iiint \frac{df}{dt} \ln f \, d\omega \quad (6-11)$$

where $d\omega$ is the range of momenta, $dp_x dp_y dp_z$. In order to determine the sign of dH/dt , we must attempt a description of the collision process and begin by rewriting Eq. (6-2).

$$n_i = f_i \delta v_i \delta \omega_i \quad (6-12)$$

We have assumed f to be independent of position and may sum over all volume elements, δv_i , to obtain the number of particles, N_i , in the container that have momentum in the range $\delta \omega_i$.

$$N_i = V f_i \delta \omega_i \quad (6-13)$$

We now consider a collision between two particles with initial momenta in the ranges $\delta \omega_i$ and $\delta \omega_j$ to produce particles with momenta ranges $\delta \omega_k$ and $\delta \omega_l$.

This type of collision, while conserving momentum, will decrease N_i and N_j and increase N_k and N_l and will occur with a frequency proportional to the product $N_i N_j$. Thus, from Eq. (6-13) we write the frequency, Z_{ij} , as

$$Z_{ij} = C f_i f_j \quad (6-14)$$

where C is the collision constant. From Z_{ij} we can determine the rate of change of the number of particles in these momentum ranges

$$C f_i f_j = -\frac{dN_i}{dt} = -\frac{dN_j}{dt} = \frac{dN_k}{dt} = \frac{dN_l}{dt} \quad (6-15)$$

Using Eq. (6-13) to evaluate these derivatives we obtain

$$\begin{aligned} -\frac{dN_i}{dt} &= -V \frac{df_i}{dt} \delta \omega_i = C f_i f_j \\ -\frac{dN_j}{dt} &= -V \frac{df_j}{dt} \delta \omega_j = C f_i f_j \\ \frac{dN_k}{dt} &= V \frac{df_k}{dt} \delta \omega_k = C f_i f_j \\ \frac{dN_l}{dt} &= V \frac{df_l}{dt} \delta \omega_l = C f_i f_j \end{aligned} \quad (6-16)$$

We now write Eq. (6-11) as a summation

$$\frac{dH}{dt} = \sum_i \ln f_i \left(V \frac{df_i}{dt} \delta \omega_i \right) \quad (6-17)$$

Using Eq. (6-16) to replace the term in parenthesis, we find the contribution of $i j$ collisions to the summation to be

$$-C f_i f_j \ln f_i - C f_i f_j \ln f_j + C f_i f_j \ln f_k + C f_i f_j \ln f_l$$

or

$$C f_i f_j \ln \frac{f_k f_l}{f_i f_j}$$

Next, we consider the reverse collision where particles with initial momenta in the ranges $\delta \omega_k$ and $\delta \omega_l$ collide to produce particles with momenta in the ranges $\delta \omega_i$ and $\delta \omega_j$. The frequency of these collisions, Z_{kl} , is

$$Z_{kl} = C f_k f_l$$

where the collision constant C is the same as for the forward collision as can be shown by the application of Liouville's equation. By the same reasoning we can write for the contribution of reverse collisions to Eq. (6-17)

$$C f_k f_l \ln \frac{f_i f_j}{f_k f_l}$$

For this pair of forward and reverse collisions, the contribution to the summation is

$$C(f_i f_j - f_k f_l) \ln \frac{f_k f_l}{f_i f_j}$$

This expression has the form, $C(x-y)\ln y/x$, which can be shown to always be negative or zero for positive values of x and y . Because the f 's are always positive, the contribution to the summation due to a pair of forward and reverse collisions will be zero or negative. Because it can be shown for spherical particles that all possible binary collisions occur in forward and reverse pairs, the summation of Eq. (6-17) will be comprised of a large number of terms that are either negative or zero. We can then state that

$$\frac{dH}{dt} \leq 0$$