Logical Relations among the Principles of Statistical Mechanics and Thermodynamics*

JOHN S. THOMSEN
Radiation Laboratory, The Johns Hopkins University, Baltimore, Maryland
(Received May 29, 1953)

Five propositions used in thermodynamics and statistical mechanics are formulated: detailed balance, microscopic reversibility, the ergodic hypothesis, the Second Law, and a restriction on the transition probabilities which will be termed the "λ hypothesis." It is shown that the last three propositions are equivalent. The ergodic hypothesis and detailed balance together are proved to be equivalent to microscopic reversibility. This shows that microscopic reversibility is a stronger condition that the Second Law. The results are discussed in connection with Onsager's principle.

INTRODUCTION

THERMODYNAMICS and statistical mechanics have been developed from a number of different starting points. Classical thermodynamics is a theory of equilibrium based on the First and Second Laws, while statistical theories of equilibrium have depended largely on the ergodic hypothesis. Statistical theories of irreversible processes, on the other hand, have generally made use of the principle of microscopic reversibility in some form. The macroscopic theory of irreversible processes also depends to a great extent on this principle, which is the basis of Onsager's reciprocal relations. Examples are found in the work of Cox,[1] de Groot,[2] Denbigh,[3] Onsager,[4] and Prigogine.[5]

At first glance these assumptions appear quite different, but a number of logical relations between them can be demonstrated. In the present paper, five basic propositions employed in statistical mechanics and thermodynamics will be formulated, and all possible relations between them will be investigated. Some of the resulting theorems are rather obvious when properly formulated, and several have been previously proved. However, all are included for the sake of a complete but concise summary of the possible relations.

I. DEFINITIONS

The following definitions will apply.

(1) *System.*—The system under consideration will be taken as one of an ensemble of N identical systems. Each one is considered isolated and each has the same energy E.

(2) *States.*—Each system is assumed to have W possible quantized states accessible to it. Since the energy is fixed, W may be taken as finite.

(3) *Probabilities.*—The probability that a system chosen at random will be in state i will be denoted by p_i where $1\leq i\leq W$. The probabilities must, of course, satisfy the restriction:

$$\sum_i p_i=1. \tag{1}$$

(4) *Transition Probabilities.*—λ_{ij} will denote the conditional transition probability per unit time of a system going from state i to state j, i.e., $p_i\lambda_{ij}dt$ is the probability of a system originally being in state i and going to state j within the small time interval dt. λ_{ii} will be defined as zero for all i. The λ_{ij}, taken collectively, may be considered as a transition matrix. All λ_{ij} are assumed to be *independent of time.*

From this definition it follows that the probability of some transition out of state i in time dt is found by summing $p_i\lambda_{ij}dt$ over all j. Similarly the probability of a transition into state i is obtained by summing $p_j\lambda_{ji}dt$ over all j. Thus the net change in p_i is

$$dp_i=\sum_j p_j\lambda_{ji}dt-\sum_j p_i\lambda_{ij}dt.$$

Hence

$$\dot{p}_i=\sum_j(p_j\lambda_{ji}-p_i\lambda_{ij}). \tag{2}$$

(5) *Equilibrium.*—The necessary and sufficient condition for equilibrium is that all $\dot{p}_i$ vanish. By Eq. (2) it follows that

$$\sum_j(p_j\lambda_{ji}-p_i\lambda_{ij})=0 \text{ at equilibrium,} \tag{3}$$

for every i.

(6) *Entropy.*—The usual statistical definition of entropy[6] will be used, namely,

$$S=-k\sum_i p_i \ln p_i, \tag{4}$$

where k is the Boltzmann constant. It might be possible to argue that this definition applies at equilibrium, but does not properly describe irreversible phenomena. However, this possibility will not be explored further, and Eq. (4) will henceforth be considered axiomatic. An expression for $\dot{S}$ may easily be obtained from Eq. (4).

$$\dot{S}=-k\sum_i\dot{p}_i \ln p_i-k\sum_i\dot{p}_i.$$

* Based on a thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at The Johns Hopkins University.

[1] R. T. Cox, Revs. Modern Phys. **22**, 238 (1950).

[2] S. R. de Groot, *Thermodynamics of Irreversible Processes* (North Holland Publishing Company, Amsterdam; Interscience Publishing Company, New York, 1951).

[3] K. G. Denbigh, *The Thermodynamics of the Steady State* (Methuen and Company, Ltd., London, 1951).

[4] L. Onsager, Phys. Rev. **37**, 405 (1931); **38**, 2265 (1931).

[5] I. Prigogine, *Étude thermodynamique des phenomenes irreversibles* (Desoer, Liege, Belgium, 1947).

[6] See, for example, R. T. Tolman, *The Principles of Statistical Mechanics* (Oxford University Press, London, 1938), Sec. 122, for a discussion of this definition.

The last term vanishes, since Eq. (1) holds identically for all time. Thus

$$\dot{S}=-k\sum_i\dot{p}_i\ln p_i. \tag{5}$$

(7) *Interconnection of States.*[7]—Two states will be said to be interconnected if it is possible to go from one to the other in both directions. (This does not mean that all λ_{ij} are non zero. For example, although λ_{ij} may vanish, it will be possible to go from i to j if $\lambda_{ik}\neq 0$ and $\lambda_{kj}\neq 0$.) In the systems to be considered it will be assumed that all states are interconnected. Mathematically this is equivalent to saying that the states of each system form an irreducible Markov chain.

II. PROPOSITIONS

Five logical propositions will now be stated; these will be denoted by letters to distinguish them from ordinary equations. It should be noted that there are some discrepancies among various authors as to the meanings of the terms used, e.g., microscopic reversibility.

(1) *Microscopic Reversibility.*—(**M**) states that the transition probabilities between two states are the same in either direction, i.e.,

$$\lambda_{ij}=\lambda_{ji} \tag{M}$$

for every i and j. This means that the transition matrix is symmetric. The definition is the same as that employed by Cox.[1]

(2) *Detailed Balance.*—(**D**) requires that transitions between any two states take place with equal frequency in either direction at equilibrium. Thus, for every i and j,

$$p_i\lambda_{ij}=p_j\lambda_{ji} \text{ at equilibrium.} \tag{D}$$

This principle is also referred to as microscopic reversibility by some authors, but in the present paper it is important to distinguish it from (**M**) as given above.

(3) *The Ergodic Hypothesis.*—(**E**) assumes that all states are equally probable at equilibrium, i.e., for every i,

$$p_i=1/W \text{ at equilibrium.} \tag{E}$$

(4) *The Second Law of Thermodynamics.*—(**S**) states that the entropy of an isolated system never decreases, i.e.,

$$\dot{S}\geq 0. \tag{S}$$

(5) *The λ Hypothesis.*—(**L**) assumes that the sum of all elements in a given row of the transition matrix is equal to the sum of the elements in the corresponding column.[8] This means that, for every i and j,

$$\sum_j\lambda_{ij}=\sum_j\lambda_{ji}. \tag{L}$$

Two simple lemmas, to be denoted by (**L$_1$**) and (**L$_2$**), will now be proved. These follow directly from (**L**) and may be used wherever (**L**) is assumed. Let f_i be any factor or product of factors which depends only on the index i. Then, by (**L**),

$$\sum_j f_i\lambda_{ij}=f_i\sum_j\lambda_{ij}=f_i\sum_j\lambda_{ji}.$$

Hence

$$\sum_j f_i\lambda_{ij}=\sum_j f_i\lambda_{ji}. \tag{L$_1$}$$

Consider next a double summation and apply (**L$_1$**).

$$\sum_i\sum_j f_i\lambda_{ij}=\sum_i\sum_j f_i\lambda_{ji}.$$

Interchanging the summation indices on the right gives

$$\sum_i\sum_j f_i\lambda_{ij}=\sum_i\sum_j f_j\lambda_{ij}. \tag{L$_2$}$$

The above propositions appear quite different at first glance. They may be placed in three categories. (**D**) and (**E**) give symmetry conditions involving the equilibrium probabilities; (**L**) and (**M**) imply certain symmetry relationships satisfied by the transition matrix; (**S**) is a thermodynamic statement of irreversibility.

III. THEOREMS

The various possible logical relationships between the above propositions will now be considered. It will first be shown that (**E**), (**L**), and (**S**) are equivalent propositions. It will then be proved that (**M**) is equivalent to (**D**) and (**E**) taken together. Finally, counter examples will be given to show that neither (**D**) nor (**E**) by itself implies (**M**). These relations may be represented schematically as follows:

$$(\mathbf{E})\rightleftarrows(\mathbf{L})\rightleftarrows(\mathbf{S}),$$
$$(\mathbf{M})\rightleftarrows[(\mathbf{E})+(\mathbf{D})].$$

Theorem 1: (**E**) implies (**L**). At equilibrium there are W equations of the form of Eq. (3), i.e.,

$$\sum_j(p_j\lambda_{ji}-p_i\lambda_{ij})=0.$$

However, (**E**) requires that all probabilities are equal to $1/W$ at equilibrium. Hence

$$(1/W)\sum_j(\lambda_{ji}-\lambda_{ij})=0.$$

It follows that

$$\sum_j\lambda_{ji}=\sum_j\lambda_{ij}. \tag{L}$$

This proves the theorem.

Theorem 2: (**L**) implies (**S**). An extreme value of $\dot{S}$ will first be found subject to the normalization condition of Eq. (1). This will be done by the method of Lagrange multipliers, which gives W equations of the form

$$\frac{\partial\dot{S}}{\partial p_r}+\mu\frac{\partial}{\partial p_r}(\sum_i p_i-1)=0.$$

[7] It will be noted that this restriction is used only in Theorems 2, 3, and 5; actually it is not essential in Theorem 2.

[8] See W. Feller, *An Introduction to Probability and its Applications* (John Wiley and Son, New York, 1950), p. 327 for a discussion of doubly stochastic matrices, which are closely related to (**L**). Let A be a matrix which gives the probabilities at time $t+\Delta t$ in terms of those at t, i.e., $p_i(t+\Delta t)=\sum_j A_{ij}p_j(t)$. For small Δt, it may be shown that $A_{ij}=\lambda_{ji}\Delta t+\delta_{ij}(1-\sum_k\lambda_{ik}\Delta t)$. If (**L**) holds, then $\sum_j A_{ij}=1=\sum_j A_{ji}$, which is the definition of a doubly stochastic matrix. The converse holds also under appropriate conditions.

With the help of Eq. (5), this becomes

$$-k\sum_i \frac{\partial \dot{p}_i}{\partial p_r}\ln p_i - \frac{k\dot{p}_r}{p_r}+\mu=0. \tag{6}$$

Since $\dot{p}_i$ is given by Eq. (2), it follows that

$$-k\sum_i(\lambda_{ri}-\delta_{ri}\sum_j\lambda_{ij})\ln p_i - (k/p_r)\sum_j(p_j\lambda_{jr}-p_r\lambda_{rj})+\mu=0. \tag{7}$$

Rearranging the first term and applying ($\mathbf{L}_1$) to the second gives

$$-k\sum_i\lambda_{ri}\ln p_i + k\sum_j\lambda_{rj}\ln p_r - (k/p_r)\sum_j(p_j\lambda_{jr}-p_r\lambda_{jr})+\mu=0.$$

When the first summation index is changed to j, μ is given by

$$\mu=k\sum_j\left[\lambda_{rj}\ln\left(\frac{p_j}{p_r}\right)+\left(\frac{p_j}{p_r}-1\right)\lambda_{jr}\right]. \tag{8}$$

Assume for the moment that all probabilities are not equal when $\dot{S}$ attains its extreme value. Let P be the probability of the most probable state or states. Since all states are assumed to be interconnected, there must be at least one state with this probability connected to a state of lesser probability. In other words, there must be an r such that $p_r=P$, $\lambda_{rj}\neq 0$ for at least one j with $p_j<P$. When this p_r is substituted in Eq. (8), $(p_j/p_r)\leq 1$ for all j, and thus μ is clearly nonpositive. However, there is at least one j described above for which $\lambda_{rj}\ln(p_j/p_r)<0$. Thus $\mu<0$.

The state or states of minimum probability may be treated similarly. At least one p_r can be found which gives $\mu>0$. Thus a contradiction is reached unless all probabilities are equal to $1/W$ at the extreme value. In this case Eq. (8) gives $\mu=0$ for every r, and the Lagrange condition is satisfied. Hence at most one extreme value exists.

When all probabilities are equal and (**L**) holds, Eq. (2) shows that all $\dot{p}_i$ vanish. It is then clear from Eq. (5) that $\dot{S}=0$; it only remains to show that this value of $\dot{S}$ is a minimum.

To investigate the region near equilibrium let the probabilities be given by

$$p_i=(1/W)(1+\Delta_i), \tag{9}$$

with the provision that $|\Delta_i|_{\max}\ll 1$. Then $\ln p_i$ may be approximated by

$$\ln p_i\approx -\ln W+\Delta_i.$$

Substitution in Eq. (5) gives

$$\dot{S}=-k\sum_i\dot{p}_i[-\ln W+\Delta_i].$$

Since $\ln W$ is independent of i and $\sum_i\dot{p}_i=0$ as a consequence of Eq. (1), this reduces to

$$\dot{S}=-k\sum_i\dot{p}_i\Delta_i. \tag{10}$$

Successive use of ($\mathbf{L}_1$) and Eq. (9) in Eq. (2) yields

$$\dot{p}_i=\sum_j(p_j-p_i)\lambda_{ji}=(1/W)\sum_j(\Delta_j-\Delta_i)\lambda_{ji}.$$

This expression is now substituted in Eq. (10); thus

$$\dot{S}=-(k/W)\sum_i\sum_j(\Delta_j\Delta_i-\Delta_i^2)\lambda_{ji}.$$

Application of ($\mathbf{L}_2$) to the second term yields

$$\dot{S}=-(k/W)\sum_i\sum_j(\Delta_j\Delta_i-\Delta_j^2)\lambda_{ji}.$$

The last two equations are now added to obtain

$$\dot{S}=(k/2W)\sum_i\sum_j(\Delta_i-\Delta_j)^2\lambda_{ji}\geq 0. \tag{11}$$

Consequently the only extreme value is a minimum and $\dot{S}=0$ at this point. Thus the Second Law (**S**) results.[9]

Theorem 3: (**S**) implies (**E**). Equation (5) shows that $\dot{S}=0$ at equilibrium since all $\dot{p}_i$ vanish. It follows from (**S**) that $\dot{S}$ must have an extreme value at equilibrium. The method of Lagrange multipliers may again be used; Eq. (6) applies since it was obtained without making any use of (**L**). The extreme value occurs at equilibrium; hence all $\dot{p}_r$ vanish and it follows that

$$-k\sum_i\frac{\partial\dot{p}_i}{\partial p_r}\ln p_i+\mu=0.$$

This can be evaluated as in the last theorem, yielding

$$\mu=k\sum_j\lambda_{rj}\ln\left(\frac{p_j}{p_r}\right). \tag{12}$$

Assume now the probabilities are not all equal, and let r be a state of maximum probability. Then it is evident that $\mu\leq 0$. By the same argument previously used, based on the interconnection of states, it is clear that at least one p_r can be found which gives $\mu<0$. Similarly, there is at least one p_r for a state of minimum probability for which $\mu>0$. Thus a contradiction is reached unless all probabilities are equal and $\mu=0$. Thus all $p_r=1/W$ at equilibrium and (E) holds.

Corollary: Propositions (**E**), (**L**), and (**S**) are equivalent.

By Theorems 1–3, (**E**) implies (**L**), (**L**) implies (**S**), and (**S**) implies (**E**). It follows that any one of them implies any other; thus they are completely equivalent.[10]

[9] A more general form of this theorem, using the concept of a doubly stochastic matrix, is given in Hardy, Littlewood, and Pólya, *Inequalities* (Cambridge University Press, Cambridge, 1952), pp. 88–91; Messenger of Mathematics **58**, 145 (1929).

[10] The equivalence of (**E**) and (**L**) is also proved in Feller (reference 4) p. 327 and in M. Fréchet, *Traité du calcul des probabilités et ses applications*, Émile Borel, Tome I, Fascicule III, Second Livre, *Théorie des événements en chaine dans le cas d'un nombre fini d'états possibles* (Gauthier-Villars, Paris, 1938), p. 37. It is obvious that (**E**) furnishes a possible equilibrium solution if (**L**) holds. However, the uniqueness of this solution involves a longer argument based on the interconnection of states. Since Theorems 2 and 3 furnish an indirect proof, the direct one will be omitted.

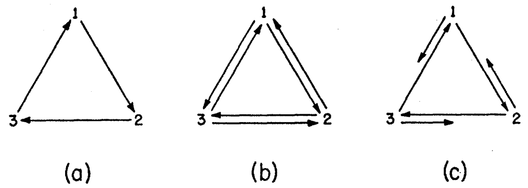

FIG. 1. Types of equilibrium for a system with three states) (Length of arrows represents number of transitions per unit time.. (a) Equilibrium given by Eq. (14), a special case of (**L**). (b) Equilibrium required by (**M**). (c) General type of equilibrium permitted by (**L**).

Theorem 4: (**E**) and (**D**) together imply (**M**). (**D**) states that $p_i\lambda_{ij}=p_j\lambda_{ji}$ at equilibrium. However, the probabilities are equal at equilibrium by (**E**); hence $\lambda_{ij}=\lambda_{ji}$ and (**M**) holds.

Theorem 5: (**M**) implies (**E**) and (**D**). (**M**) is clearly a special case of (**L**). However, (**L**) is equivalent to (**E**); hence (**M**) implies (**E**). Since all probabilities are thus equal at equilibrium and $\lambda_{ij}=\lambda_{ji}$ by (**M**), it follows that

$$p_i\lambda_{ij}=p_j\lambda_{ji} \text{ at equilibrium.} \qquad (\mathbf{D})$$

This proves the theorem.

The relations stated at the beginning of this section have now been proved. It might still be thought that there are other relations which have been overlooked. Such a possibility will be disproved by means of counter examples in the two succeeding theorems.

Theorem 6: (**D**) does not imply (**M**). Let $p_1=\frac{1}{2}$, $p_2=\frac{1}{3}$, and $p_3=\frac{1}{6}$, and let the transition matrix be

$$\lambda=\begin{bmatrix}0 & 2 & 1\\ 3 & 0 & 1\\ 3 & 2 & 0\end{bmatrix}. \qquad (13)$$

Equation (3) is satisfied, showing that the distribution represents a condition of equilibrium. (**D**) is seen to hold, but (**M**) is obviously not satisfied. Actually it will be possible to find a λ matrix which satisfies (**D**) for any given set of equilibrium probabilities; (**M**) will not hold true unless the probabilities in this set are all equal.

Theorem 7: (**L**) does not imply (**M**). Consider the transition matrix:

$$\lambda=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{bmatrix}. \qquad (14)$$

Each row and column adds to unity, thus satisfying (**L**). However, obviously (**M**) does not hold.

CONCLUSIONS

It is interesting to compare the types of equilibrium given by (**M**) and (**L**) respectively. It has been shown that (**E**) holds in both cases. If the matrix of Eq. (14) is taken as an example of (**L**), the equilibrium picture will be that indicated in Fig. 1(a), with all transitions occurring in the same direction. This equilibrium, which may be termed "circular," has been discussed by Onsager[4] and Denbigh.[11]

A symmetric matrix, i.e., one which satisfies (**M**), will give detailed balance (**D**), as proved in Theorem 5. The equilibrium situation for three states is shown in Fig. 1(b). If a symmetric λ matrix is added to that of Eq. (14), (**L**) is still satisfied by the sum. This gives the most general type of equilibrium consistent with (**L**) and is represented in Fig. 1(c).

Obviously (**M**) is a stronger requirement than (**L**) or (**S**) and may thus yield valuable information not obtainable from classical thermodynamics. Microscopic reversibility is frequently referred to as an auxiliary to the First and Second Laws.[12] However, when (**M**) is stated in the form given here, it includes (**S**) as a consequence, without involving any additional assumption except the statistical definition of entropy.[13] Clearly, it is also possible to postulate (**S**) and (**D**) and reach the same conclusions as obtained by assuming (**M**).

There is one disconcerting question raised by these results. It might appear, on the basis of quantum mechanics, that (**M**) should hold under all circumstances.[14] (Since the definitions employed are not necessarily valid for all quantum-mechanical problems, this conclusion may not always be true.) According to Onsager,[5] microscopic reversibility as he defines it cannot be expected to hold in the presence of a magnetic or Coriolis field. It is not obvious just how this feature of Onsager's principle is to be reconciled with quantum mechanics. This point may be discussed further in a future paper.

The author wishes to thank Professor Richard T. Cox of The Johns Hopkins University for suggesting the problem discussed here and for giving many hours of help toward the preparation of this paper. He is also indebted to Professor Theodore H. Berlin, Dr. Albert J. Novikoff, and Mr. David B. Lowdenslager, all of The Johns Hopkins University, for advice and assistance.

[11] Reference 3, pp. 31–34.
[12] See, for example, reference 3, p. 96.
[13] See Cox, reference 1, for a more direct proof of this statement.
[14] See, for example, P. A. M. Dirac, *Quantum Mechanics* (Oxford University Press, London, 1947), third edition, pp. 172–74.