

equally tenable to those involved in the microscopic derivation since both are consistent with empirical observations on related physical phenomena.

The principal result of the Onsager development is that the reciprocal relations, derived by application of the theorem of microscopic reversibility, permit a direct comparison of fluxes and forces with physically, identifiable quantities. On the other hand, the macroscopic derivation presented herein achieves the same result by virtue of the fact that dW and $d\dot{W}$ can be treated as exact differentials for the conditions under which the equations of irreversible thermodynamics hold and to the extent that the fundamental laws of classical macroscopic thermodynamics are valid. In other words, since the lost work is already known a priori to be path independent (when no work is done at any stage of the process) no new information is gained by resorting to the theorem of microscopic reversibility.

ACKNOWLEDGMENT

The criticisms of F. Andrews and F. Mixon were invaluable. This work was supported in part by the Air Force Office of Scientific Research, Grant AF-AFOSR-563-65.

REFERENCES

1. Andrews, F. C., *Ind. and Eng. Chem. Fund.* 6, 48, 1967.
2. Bearman, R. J. and Crawford, J. G., *J. Chem. Phys.* 28, 136, 1958.
3. Coleman, B. D. and Truesdell, C., *J. Chem. Phys.* 33, 28, 1960.
4. deGroot, S. R. and Mazur, P., *Non Equilibrium Thermodynamics*, Chapter XV, Interscience Publishers, Inc., New York, 1962.
5. Kirkwood, J. G. and Crawford, B., Jr., *J. Phys. Chem.* 56, 1048, 1952.
6. Martin, J. J. "The Symmetrical Fundamental Property Relations of Thermodynamics," Presented at the San Francisco meeting of the American Institute of Chemical Engineers, May 1965.
7. Onsager, L., *Phys. Rev.* 37, 405, 1931; 38, 2265, 1931.
8. Sliepcevich, C. M. and Finn, D. in *Chemical Engineers' Handbook*, 4th ed., pp. 4-42, 4-44, and 4-69, McGraw-Hill Book Co., Inc., New York, 1963.
9. Sliepcevich, C. M. and Finn, D., *Ind. Eng. Chem. Fund.* 2, 249, 1963.
10. Sliepcevich, C. M., Finn, D., Hashemi, H., and Heymann, M., *Ind. Eng. Chem. Fund.* 3, 276, 1964.
11. Sliepcevich, C. M. and Hashemi, H. T. "Reciprocal Relations in Irreversible Processes." Presented at the Philadelphia meeting of the American Institute of Chemical Engineers, December 1965.
12. Wei, James, *Ind. and Eng. Chem.* 58, 55, 1966.

SUMMER, 1968

APPROACHES TO STATISTICAL THERMODYNAMICS*

M. V. SUSSMAN

*Tufts University
Medford, Mass.*



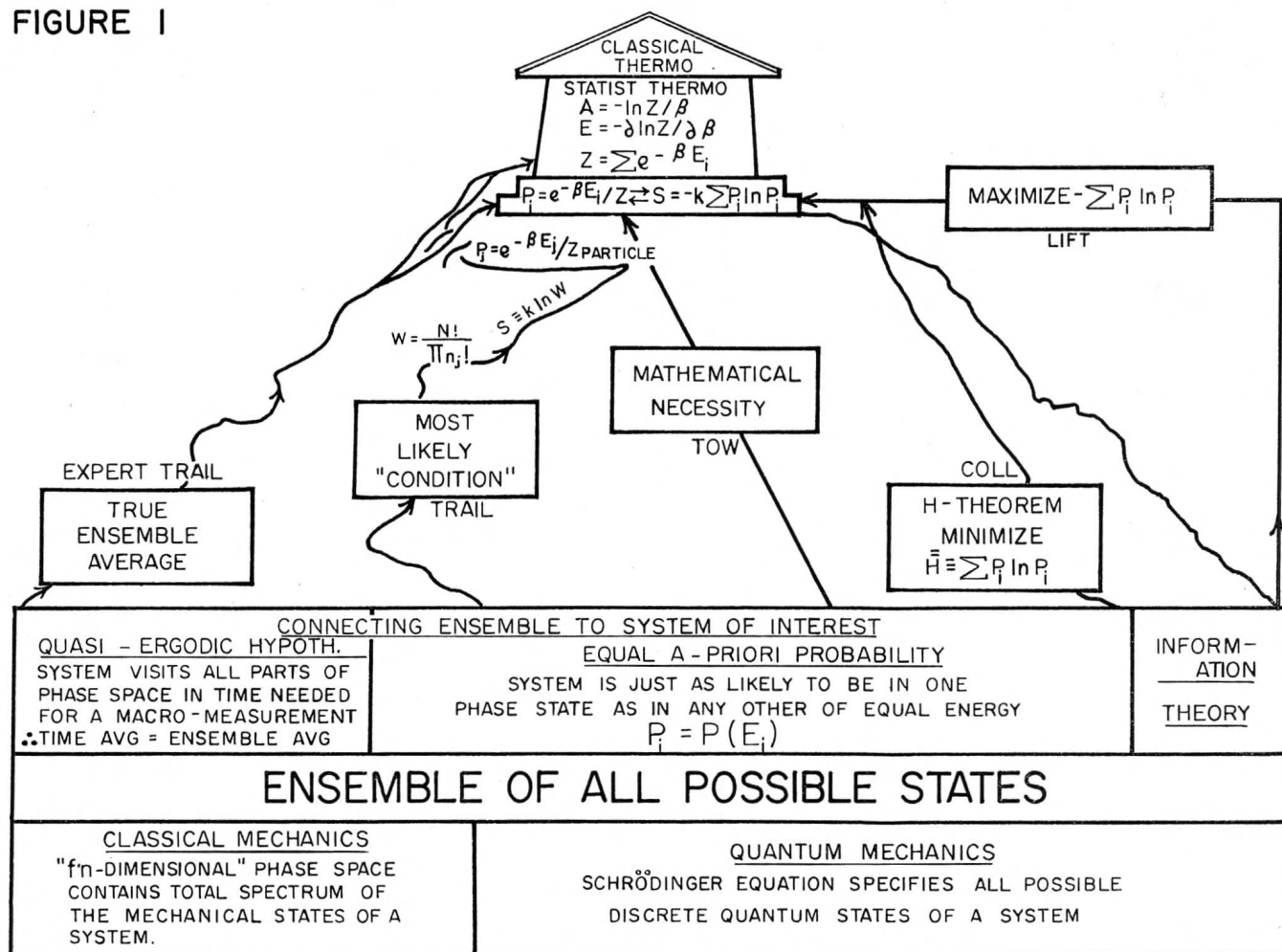
Statistical thermodynamics connects classical thermodynamics which describes the energetic interactions of macroscopic systems with the properties of the microscopic or molecular constituents of a system. The connection expands the application of thermodynamics to extreme temperature, solid state, thermo-electric, and other phenomena. It permits derivation of equations of state, and calculation of thermodynamic properties from spectroscopic data. It provides insights to many thermodynamic properties, particularly the entropy. Like many other worthwhile goals, statistical thermodynamics may be approached in a number of ways. The various approaches each have their strong proponents and detractors and the selection of an approach is often a subjective decision reflecting the user's mathematical sophistication, epistemological philosophy and teacher's prejudice.

My purpose here is to outline the more common approaches to statistical thermodynamics, necessarily in qualitative terms and with more emphasis on the similarities than the differences. My point of view is summarized by the mountain-scape sketched in Fig. 1. In the brief time available I will run you over the various trails, passes and pathways which have been used to connect microscopic to macroscopic thermodynamic behavior.

*Presented at the Annual Meeting of ASEE, June 19-22, 1967.

113

FIGURE 1



Statement of Basic Problem

All approaches to the problem start from the following common ground.

1. Recognition that every macroscopic system has a fantastically detailed microscopic structure, and that the existence of this micro structure makes possible an astronomically large number of different arrangements of the microscopic elements (quantum states) which are completely consistent with the macroscopic system's properties.
2. A realization that there is no way of knowing which arrangement or state actually represents the system and therefore, all (or a most representative portion of) the possible micro-states must be considered in determining the system's properties.

The basic problem of statistical thermodynamics is therefore the assignment of a weight (a probability) to each possible micro-state which reflects its contribution to the properties of the macroscopic system.

It is in the rationalization of the averaging technique, that is, in the derivation of the func-

tion (called a "distribution" function) assigning weight or probability to each micro state, that a variety of approaches are used. All approaches arrive at essentially the same result: For a closed constant-volume system in equilibrium with a heat bath the probability of the i'th micro-state is equal to

$$P_i = \frac{1}{Z} \exp - \beta E_i \tag{1}$$

where E_i is the energy of state i and β and Z are constants of the equilibrium system.

The sum of all the probabilities = 1, and therefore

$$\sum P_i = 1 = \sum_i \frac{1}{Z} \exp - \beta E_i$$

$$Z = \sum_i \exp - \beta E_i \tag{2}$$

Z is called the "partition function," or "sum over states."

β is shown to be $1/kT$. The expected energy of the macroscopic system is equal to:

$$\langle E \rangle = - \frac{\partial \ln Z}{\partial \beta} \quad (3)$$

and the entropy of the system is equal to

$$S = - k \sum P_i \ln P_i \quad (4)$$

$$(\text{or } S = k \ln W) \quad (5)$$

From here the expressions of classical thermodynamics are obtained by straightforward, unsophisticated mathematical techniques.

Ensemble of States

Let us now explore the Fig. 1 mountain beginning at its base—the concept of an ensemble of all possible microscopic states of a macroscopic system. In quantum mechanics, the Schrödinger equation specifies the possible discrete macroscopic or quantum states of a system. The totality of these states is the quantum mechanical representation of the ensemble. An alternate and older view of the ensemble is provided by classical mechanics where a many dimensional hyperspace is used to chart the total spectrum of mechanical states of all the microscopic constituents of the system that are consistent with the macroscopic knowledge about the system. This hyperspace is called the “phase space” of the system.

Having set up the ensemble of all possible states in either quantum mechanical or classical mechanical terms, it becomes necessary to connect the ensemble to the macroscopic system of interest. The connection is made in the ways indicated in Fig. 1.

Quasi-Ergodic Hypothesis

The average properties of an ensemble are related to the properties of a given macroscopic system by making an assumption about the actual *mechanical* behavior of the macroscopic system, viz:

A property measurement (for example pressure) made on a macroscopic system is a time average property measurement rather than an instantaneous property measurement. The measurement time is long on a microscopic scale and within the measurement time interval the system visits (or comes arbitrarily close to) all points in the phase space of the ensemble. It therefore follows that a time average property of a macro-system is the same as an ensemble average property.

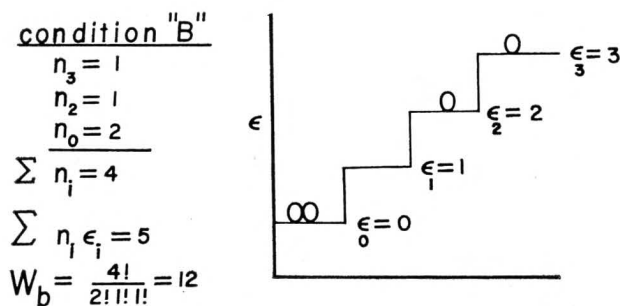
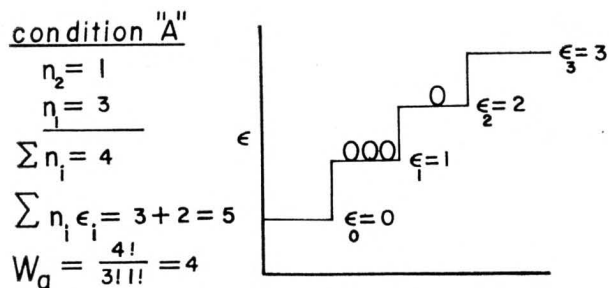


Figure 2.—Most likely “condition.” Condition “B” is more likely than condition “A” because $W_b > W_a$.

The validity of the ergodic hypothesis is questionable particularly because systems can be imagined where the hypothesis does not hold; for example, an ideal gas in a rigid parallel wall container whose particles are so arranged as to move perpendicular to the parallel faces of the container, and in such a manner that no collision occurs between the particles. This system would not visit all regions of phase space, that is go through all configurations of its particles’ positions and velocities consistent with the total energy of the system.

Equal A-Priori Probabilities

Another method of connecting the ensemble to the macro-system of interest is to assign equal statistical weight or probability to all equal micro-states of the ensemble. This is a reasonable assumption because knowing only the energy of the system, we have no basis for choosing one micro-state over any other micro-state having the same energy. The system has an equal likelihood of being in all such micro-states. Therefore, its *average* property is the average over all the equally likely states.

A corollary of this approach is that the probability of a micro-state is a function of the energy of that state only, that is,

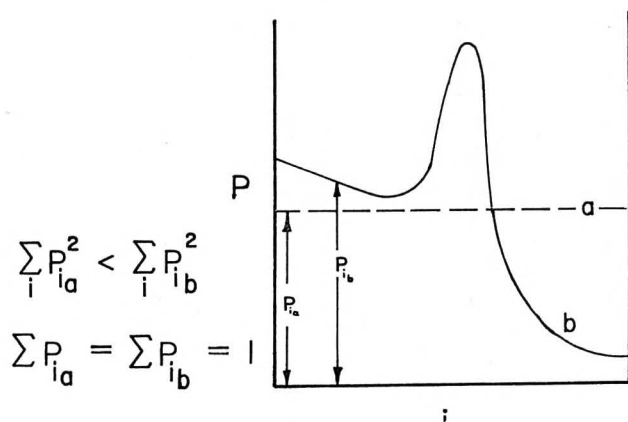


Figure 3.—Moment of a distribution.

$$P_i = f(\text{Energy of state } i) \quad (7)$$

The third way of connecting an ensemble to the system of interest indicated in Fig. 1 is the information theory approach which implicitly agrees with the equal probability assumption, although it does not make the assumption explicitly. More will be said about this later.

True Ensemble Average

We now turn to the trails ascending from our base camp to the "distribution function" (Fig. 1).

Given that ensemble average properties are the same as the macroscopic properties of a system, the system property (M) is found by integrating over phase space

$$M = \int \rho(p, q, t) M(p, q) dpdq \quad (8)$$

where p and q are the generalized coordinates of phase space; t is time and ρ is a density function which gives the probability of finding a state point in any unit volume of phase space. A mathematical theorem due to Liouville is then used to show that the density function is independent of time

$$d\rho/dt = 0$$

if ρ is a constant or a function of the energy of the entire system. (When this condition prevails the ensemble is said to be in statistical equilibrium). A suitable function is $\rho = \exp -(\lambda + \beta E)$ which leads to the conclusion that the probability of a state is proportional to exponential $(-\beta E_{\text{state}})$.

This is the route taken by the professional statistical mechanic. It requires considerable mathematical sophistication. It is thorough, elegant, rigorous, and generally unsuitable for presenting the useful concepts of statistical thermodynamics to undergraduates.

Most likely "Condition"

An alternate route to the "distribution function" I have called the "most likely condition." It is supposed to be a short cut since it attempts to evaluate the average property of an ensemble, not by covering all states in the ensemble, but only the most likely states, as represented by the most likely "condition." The "condition" of a system is the set of occupancy numbers (n_i) which designate the number of microscopic particles in each of the energy levels accessible to a system's particles. For example, Fig. 2 shows a system which has only four particles. The "a-condition" of that system is given by the set of occupancy numbers (n_i); $n_1 = 3$; $n_2 = 1$. The sum of the n_i is equal to the total number of particles in the system, in this case 4; and the energy of the system is equal to

$$E = \sum n_i \epsilon_i = (3 \times 1) + (1 \times 2) = 5 \text{ energy units}$$

Now, three 1-energy unit particles and one 2-energy unit particle can be permuted in $4!/3!1! = 4$ ways. (The general rule for the number of permutations of N total objects where N is equal to $\sum n_i$; is $W = N!/\pi n_i!$) Condition "b", given by $n_3 = 1$, $n_2 = 1$, $n_0 = 2$, allows for 12 accommodations or permutations. Therefore, if we were betting on condition "a" or "b" we would put our money on "b" as the more likely "condition."

Quite clearly, the most likely "condition" of any system is that set of n_i 's (consistent with the system's energy) which produces the maximum number of permutations. It can be shown that as the number of particles becomes very large the likelihood of any condition other than the most likely condition becomes very small. Therefore, the ensemble as a whole can be described with reasonable accuracy in terms of its most likely "condition" and the set of n_i 's that correspond to that most likely condition is simply found by maximizing the number of permutations W or $\ln W$ taking into account the fact that $\sum n_i = N$ and $\sum n_i \epsilon_i = E_i$. This technique if followed carefully, and if certain pitfalls are avoided, eventually leads to an expression for the partition function of a multiparticle system in

terms of the allowed energy levels of its constituent particles. The pitfalls and somewhat odd rationalizations* used to arrive at this final result offset the shortcut promised by averaging over the most likely condition rather than over the entire ensemble. In this approach $S \equiv k \ln W$.

Mathematical Necessity

Using the equal a-priori probability assumption, the probability of a state is a function only of its energy (see eq. 7). If we have two systems at equilibrium with a thermostatic bath whose size is such that fluctuations of the energy of one system will have no effect on the energy of the bath or the energy of the other system, then we can state

$$P_i = f(E_i) \quad (7)$$

$$P_j = f(E_j) \quad (9)$$

where E_i represents an allowed energy state of the first system and E_j represents an allowed energy state of the second system. Now, considering both systems together, the probability of the first system being at E_i and the second system at E_j must be

$$P_{i \text{ and } j} = f(E_i + E_j) = P_i P_j \quad (10)$$

$$\text{therefore } f(E_i + E_j) = f(E_i) f(E_j) \quad (11)$$

The only function satisfying (11) is an exponential

$$\text{Therefore } P_i = f(E_i) = \frac{1}{Z} \exp -\beta E_i \quad (1)$$

and we are again at the top of the mountain. A mathematical consequence of (1) and the classical definition of entropy is that S , can be shown to be equal to $S = -k \sum P_i \ln P_i$ (13) This is the approach taken by Denbigh,² Andrews,¹ and others. It is straightforward enough to be taught to undergraduates, requiring only acceptance of the fact of the existence of a multitude of quantum states and the assumption of the equal probability of equal energy quantum states. A maximization computation is avoided.

Information Theory Approach

The Information Theory approach, while using exactly the same mathematical forms established in the older statistical thermodynamic literature, has a somewhat different philosophical or logical orientation. It states that statistical

*Particles are assumed to be distinguishable. Also, Stirlings approximation, $\ln n! = n \ln n - n$, is used.

thermodynamics is not a physical theory whose validity depends either on the truth of additional basic assumptions, such as ergodic behavior or equal probability, or on experimental verification. It is instead a form of statistical inference; a technique for making the best estimates on the basis of incomplete information. If experimental verification is not obtained this is not a shortcoming of the statistical thermodynamics, but of the information supplied.

The relationship $S = -k \sum P_i \ln P_i$ (13) occupies the primal position in this approach. The equation is the basic equation of Shannon's "Mathematical Theory of Information" and is identified with thermodynamic entropy. Maximizing (13) subject to the constraints that

$\sum P_i = 1$, (The system must be in some state) and $\sum P_i E_i = E$; (The system has energy $\langle E \rangle$) leads immediately to

$$P_i = \frac{1}{Z} \exp -\beta E_i \quad (1)$$

It is the contention of the information theorists that maximizing $-\sum P \ln P$ subject to constraints produces the least biased distribution of probabilities; a distribution which is maximally non-committal with regard to missing information.

An identical technique using a different rationale was suggested by Pauli who showed that the distribution functions are obtained by minimizing the Boltzmann H- function

$$H \equiv \sum P_i \ln P_i$$

subject to constraints. The latter technique is discussed in detail by Tolman.³ Taking $\sum P \ln P$ to an extremum is not a new idea. The "information theorists" however give it new importance by insisting that it is the **most fundamental** approach to statistical mechanics, because evaluating the P_i is a problem in guessing (i.e., statistics) and not physics, and therefore there need be no further concern with Ergodicity or Equiprobability and their justification.

From an undergraduate teaching point of view, the information theory approach is almost as simple as the previously mentioned mathematical necessity approach. The student is asked to accept, without proof, the axiom that maximizing S subject to the known properties of a system produces a minimally biased set of P_i 's. The mathematics of maximization are reasonably straightforward. The trouble with the axiom is that it does not relate to much in the undergradu-

ate's experience whereas other thermodynamic and mathematical axioms usually have some intuitive acceptability.

Smoothing Function

A way of making the axiom more acceptable is to demonstrate qualitatively that maximizing $-\sum P_i \ln P_i$ or minimizing $+\sum P_i \ln P_i$ is a smoothing operation which tends to minimize the "moment" (lower the center of gravity) of a plot of P_i vs. i .

As qualitative example, assume that we have a system which is capable of existing in a great number of possible states, and we are asked to arbitrarily assign probabilities to each of these states. The states can be ordered in a sequence, and indexed by an integral subscript i . Assume that all we know about this system is that it must be in some state P_i . In Fig. 3, line b is an arbitrarily assigned distribution for this system which is constrained only by the fact that the sum of the ordinates equals unity, that is $\sum P_i = 1$. This is *not* an unbiased distribution because I have put maxima and minima in this distribution, that is, I have given some states more weight than others, without information that would justify so doing. The relative smoothness of the arbitrary curve in Fig. 3, can be represented by the mathematical index

$$\sum P_i^2 \quad (14)$$

which evaluates the "moment" of the distribution about the horizontal axis. The "moment" increases as the magnitude of the singularities or extrema in the system increase, and conversely, decreases as the center of gravity of the distribution drops, that is, as the curve becomes more uniformly *smooth*. In fact, it is a straightforward exercise in calculus of variation to show that the minimum "moment" corresponds to line "a," a constant value of P_i which is certainly the smoothest possible curve. If, in the smoothness index, (Eq. 14) we replace one P_i with a monotonic function of P_i , that is $-\ln P_i$, we should expect similar behavior. In other words, the effect of maximizing $-\sum P_i \ln P_i$ is to *smooth* out our distribution. The advantage of the logarithmic function is that it allows expressing S as a function of the probability of the microstates, and it prevents P_i from taking on negative values.

Allow me to end with a speculative aside. Maximizing entropy smooths a distribution. This

suggests to me that it might be possible to restate the principle in terms of geodesics. I say this because I would assume that a maximally smooth curve should have a minimum arc length.

I have tried using a criterion of "minimum arc length" to find a distribution function, harboring secret hopes that the criterion would lead to $-\sum P_i \ln P_i$ and even more general expressions for new entropies. I regret that I've not been successful. The geodesic idea (that an unbiased distribution has minimum arc length) nevertheless continues to intrigue me and I would welcome thoughts of others on how to work it into a selection formalism.

References

1. Andrews, F. C., "Equilibrium Statistical Mechanics," John Wiley & Sons, Inc., 1963.
2. Denbigh, K., "The Principles of Chemical Equilibrium," Cambridge University Press, 1961.
3. Tolman, R. C., "The Principles of Statistical Mechanics," Oxford University Press, 1942.

Dr. M. V. Sussman is Professor of Chemical Engineering and Department Chairman at Tufts University and is presently on leave with NSF in New Delhi working on an Indian Engineering Education program. In August he will be an NIH Fellow at the Weizmann Institute studying biological thermodynamics and mechanochemistry. Dr. Sussman has degrees from City College of New York (BChE) and Columbia University (MS and PhD). Thermodynamics is a compulsive hobby with him and some of the ideas expressed in this article will appear in a book to be published by Wiley.

Join THE AMERICAN SOCIETY FOR
ENGINEERING EDUCATION
and Receive
CHEMICAL ENGINEERING EDUCATION

W. Leighton Collins, Executive Secretary
The American Society for Engineering Education
2100 Pennsylvania Avenue, North West
Washington, D. C. 20037

Dear Mr. Collins,

Please send me an ASEE application blank I would like to join the Chemical Engineering Division.

Name

Address