IRREVERSIBLE STATISTICAL MECHANICS AND THE PRINCIPLE OF MAXIMUM ENTROPY

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF PHYSICS
AND THE COMMITTEE ON THE GRADUATE DIVISION
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

By
Douglas James Scalapino
May 1961

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Edwin T. Jaynes

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

James H. Burgess

Approved for the University Committee
on the Graduate Division:

A H Bowker

Dean of the Graduate Division

ACKNOWLEDGMENT

I wish to thank Professor E. T. Jaynes for the many, many, discussions during which this work was formulated and carried out. I am deeply indebted to him for his continued encouragement, and many valuable suggestions. He has been a teacher and a friend. Thank you, Ed.

I am grateful to the Physics Department of Washington University for a research assistantship during the school year 1960-1961.

It is a pleasure to thank my wife, Diane, for typing this dissertation during the week before our son was born.

TABLE OF CONTENTS

CHAPTER

VI. (cont'd)

APPENDIX

LIST OF GRAPHS

CHAPTER I

INTRODUCTION

The purpose of statistical mechanics is to establish relationships between the observed macroscopic properties of systems composed of many particles, and the mechanical properties of these particles and their interactions. Statistical mechanics attempts to proceed from the laws of microscopic dynamics to the prediction of macroscopic results. These dynamic laws have the form of differential equations. Their solution depends upon the initial state of the system. The formalism of statistical mechanics must therefore provide: 1) a method for constructing a suitable representation of the initial state of the system; 2) a method of analyzing the relevant part of the many-body dynamical problem.*

This first aspect requires an extension of the theory of measurements. Both classical and quantum mechanics contain prescriptions for constructing initial states if the results of a complete measurement are available. For a macroscopic system such a measurement cannot be performed. Only a small number of dynamic variables are actually measured. Even for a microscopic system a complete measurement is never actually performed. Therefore, extensions to the theory of measurements have been given. This is a problem of statistical

---

*It is important to keep in mind the fact that statistical mechanics seeks to predict macroscopic results. This means that 1) and 2) need not represent the exact microscopic aspects of the initial state and its dynamic development. They need only correspond to the relevant macroscopic behavior of the system.

inference. In the quantum mechanical case*, the prescriptions generally consist of methods for constructing initial density matrices or allowed manifolds of states which are to be averaged over.

One such prescription$^{1)}$ is the method of "equal a priori probabilities and random a priori phases of accessible states." This is based upon Liouville's Theorem and Laplace's "Principle of Insufficient Reason." However, the available information is seldom of the type which permits an obvious resolution of the states into accessible and inaccessible ones. It is also difficult to generalize this method beyond the problems of thermodynamic equilibrium.

Recently, a general method has been given for the construction of density matrices based upon partial knowledge of the state of the system.$^{2)}$ This method of statistical inference is called a maximum-entropy estimate (hereafter referred to as MEE). It gives the least biased density matrix consistent with the available information. E. T. Jaynes$^{2)}$ has shown that the MEE formalism reproduces the canonical equilibrium statistical mechanics of Gibbs.

MEE is applicable to any system. In particular, it can be used to construct a density matrix for a microscopic system when the measurement is less than complete. The MEE

---

*We shall discuss the quantum mechanical situation. Similar results exist for the classical case.

1) R. C. Tolman, The Principles of Statistical Mechanics (Oxford University Press, London, 1948), Chapter IX, Sec. 84.

2) E. T. Jaynes, Phys. Rev. 106, 620 (1957); Phys. Rev. 108, 171 (1957).

formalism is the only general method so far proposed for treating this situation. A complete measurement can, in principle, be made on a microscopic system. In this case the state of the system can be determined within the framework of quantum mechanics. This limiting case provides an interesting check of the consistency of the MEE formalism. In the case of a complete measurement, the density matrix must reduce to a projection operator specifying a single "pure" quantum state. In Chapter II we investigate this problem for the case of a non-relativistic electron.

A MEE density matrix should not be identified as describing the "true" microscopic state of the system. Rather it represents in some sense the manifold of possible states consistent with the experimental observations. In order to clarify the precise meaning of this we investigate several properties of a MEE for a macroscopic spin system, Chapter III. We prove two theorems concerning a high probability manifold of states characterized by the MEE density matrix. These theorems were originally proved under different assumptions by C. E. Shannon.3) This discussion allows us to identify the high probability manifold of a MEE density matrix with the manifold of experimentally allowed states.

The statistical mechanics of a system in thermodynamic equilibrium requires only this specification of the initial state. The detailed dynamic aspects of the many-body problem,

---

3) C. E. Shannon, Bell System Tech. J. 27, 379, 623 (1948). In Appendix A we review some of the results of information theory.

part (2), are thereby avoided. More generally all equilibrium problems involve only the characterization of the initial state. By equilibrium we mean that the values of the macroscopic observables are stationary in time. This includes the class of stationary states in which external constraints (concentration gradients, temperature gradients, etc.) prevent the system from reaching thermodynamic equilibrium. In their Encyklopädie article on statistical mechanics, P. and T. Ehrenfest[4] ask whether the distribution function which describes the stationary state of an irreversible process could be considered as in some sense "relatively most probable," and whether it could be characterized as determining the extremum of some function?[5]

The MEE formalism provides an answer to this question. In Chapter IV we consider the problem of stationary, constrained equilibrium. Initially, the equation of state for a spin system

$$f(M,H,T) = 0$$

is discussed. This is a thermal equilibrium problem. If the

---

[4] P. and T. Ehrenfest, Ency. d. Math. Wissen, Bd. IV, No. 32 (Leipzig 1911), Footnote 231, p. 82.

[5] I. Prigogine has suggested that the principle of minimum entropy production is the answer to this question. Etude Thermodynamique des Phenomenes Irreversibles (Liege, 1947) Chap. V. However the actual usefulness of this principle is questionable. M. J. Klein, Termodinamica Dei Processi Irreversibili, edited by N. Zanichelli (Societa Italiana Di Fisica, Bologna, 1960) p. 203, states, "It is somewhat doubtful, however, whether the principle (of minimum entropy production) gives a characterization of the steady state which is really comparable to the characterization of the equilibrium state as that of maximum entropy. ...From both our statistical development and from the thermodynamic development of the principle one sees that the information needed to apply the principle is always large enough so that the principle supplies nothing new."

initial information consists of the temperature T and the field H, the standard computational procedure of equilibrium statistical mechanics allows the magnetization M to be calculated. Using this standard formalism of equilibrium statistical mechanics, it is not clear how to calculate T when M and H are given. However, in a consistent theory the same equation of state must emerge whether the initial information specifies (T,H), (H,M) or (M,T). The MEE formalism allows us tc treat this problem. The important point is that when H and M are given we must find an explicit manner of stating that the system is in equilibrium. We require that the entropy be maximized over the manifold of ρ's which commute with the Hamiltonian and give the specified value of M. This prescription reproduces the equilibrium equation of state.

Using this form of the equilibrium constraint a general formalism for constrained equilibrium problems is given. As in the case of thermodynamic equilibrium the problem is simply to construct a representation of the initial state. While this avoids the explicit problems associated with the dynamical time development, this stationary formulation does not remain completely separated from the dynamics. We find that the initial density matrices depend upon the "diagonal parts" of the observed operators F. The "diagonal part" of an operator F is essentially the d.c. part of the Heisenberg operator F(t). F(t) obeys the equations of motion

$$\dot{F}(t) = \frac{i}{\hbar}[\mathcal{H}, F(t)]$$

Using this constrained equilibrium formalism we investigate the problem of steady diffusion. For a small concentration gradient the diffusion tensor is calculated. We obtain a relation having the same form as the transport coefficient relations calculated by R. Kubo et. al.6) In principle our method could be used to treat, in a higher order approximation, non-linear effects occurring in constrained equilibrium problems. Kubo's theory is not applicable to these effects since he treats a system which is evolving in time toward thermal equilibrium.

In Chapters V and VI we treat certain aspects of the dynamic time development of the many-body system. In Chapter V we discuss the Second Law of thermodynamics from the point of view of the MEE formalism. The relationship between the information theory entropy and the experimental entropy is discussed. A definition of entropy for non-equilibrium states is given. The time development of this entropy is shown to obey the Second Law. The problem of the conditions under which a laboratory process is isentropic is discussed. As an example we consider a spin system in an external field which is changing in time.

In the final chapter (VI), we investigate two related problems of the dynamic behavior of a macroscopic system.

---

6) R. Kubo, M. Yokota, and S. Nakajima, Journ. Phys. Soc. (Japan), 12, 1203 (1957).

The first of these is the time decay of equilibrium correlation fluctuations. The second is the irreversible tendency of macroscopic systems to approach a state of equilibrium. These two aspects of irreversibility are related. The time dependence of a dynamic variable depends upon the correlation functions of the interaction responsible for its time development.

Two examples are given in which a correlation function can be exactly calculated. The first of these is for the ring of N exchange-coupled spins. The kth spin-up correlation function, $\langle \eta_k^+ \eta_k^+(t) \rangle$, is calculated for a particular initial state. The operator η_k^+ gives 1 if the kth spin is up and 0 if it is down. This problem provides a simple illustration of the Poincaré recurrence phenomenon. The second example is the calculation of correlation functions for thermal equilibrium electromagnetic fields (black-body correlation functions).

These examples show that the fundamental cause of irreversible decay is a dephasing, and consequent destructive interference among the terms of a sum of matrix elements. The conditions on the initial state of the system, the Hamiltonian $\mathcal{H}$ and the observed operator F under which this dephasing persists over the range of laboratory times are discussed. The long time behavior of the expectation value of a dynamic variable can be characterized by a distribution of poles $(s - iE'')^{-1}$. 7),8) This distribution is determined by the

---

7) S. Golden and H. C. Longuet-Higgins, Journ. of Chem. Phy. 33, 1479 (1960).

8) N. Saito, Phys. Rev. <u>117</u>, 1163 (1960).

initial state ψ_0 , the Hamiltonian $\mathcal{H}$, and the operator F. It is $m(E'') = \int dE' \, F(E''+E', E')$ with,

$$F(E,E') = \int d\alpha \, d\alpha' \, \langle \psi_0 | E\alpha \rangle \langle E\alpha | F | E'\alpha' \rangle \langle E'\alpha' | \psi_0 \rangle$$

where $|E\alpha\rangle$ represents an energy eigenstate of $\mathcal{H}$.

We analyze the time response for various distributions of poles. It is shown that $\langle F(t) \rangle$ decays to zero if the distribution F(E,E') has no singularities as strong as a δ-function. For $\langle F(t) \rangle$ to irreversibly approach a non-zero equilibrium value, F(E,E') must contain a $\delta(E-E')$ singularity. If F(E,E') contains singularities of the form $\delta(E-E' \pm \omega)$ the long time behavior of $\langle F(t) \rangle$ is oscillatory.

CHAPTER II

Maximum Entropy Estimates for Microscopic Systems

Quantum mechanics provides a prescription for converting the information contained in a wave function or density matrix into predictions about experimental measurements. However, the equally necessary inverse of this process, whereby one may construct a wave function or density matrix given certain experimental results is, in general, not uniquely provided within the framework of quantum mechanics. The standard approach9) to this problem assumes that if the result of measuring a dynamic variable D gives a value D' lying within $D'' < D' < D'''$ the wave function is given by,

$$\psi = \frac{\sum_n' a_n u_n}{\sum_n' |a_n|^2} \tag{2-1}$$

where $Du_n = D_n' u_n$ and the summations are only over the values of n for which $D'' < D_n' < D'''$. The coefficients a_n are unspecified. If the eigenvalues of D are non-degenerate and discrete, we may, with a sufficiently precise measurement, reduce the above summations (2-1) to a single eigenstate. In this case the standard quantum mechanics formulation does provide a unique means of constructing a wave function from experimental observations.

More generally, if the eigenstates of D are discrete, but degenerate, a sufficiently precise measurement of D reduces

---

9) F. Mandl, Quantum Mechanics (Butterworths Scientific Publications, London, 1954), Chapter II.

the summations (2-1) to the manifold of states all having the eigenvalue D'. However, within this manifold the coefficients a_n are unspecified. In order to specify these coefficients the notion of a complete set of commuting observables is introduced. That is, let L commute with D. Now a measurement of L will allow a possibility of further reducing the summations in (2-1). The sum will then extend over the intersection of the manifold of states characterized by the eigenvalues D' and L'. This process is continued until a complete specification of the state of the system is obtained. Only in this limiting case does quantum mechanics provide a unique method of translating experimental results into its mathematical formalism.

For a microscopic system the complete measurement outlined above is, at least in principle, a possibility. For macroscopic systems it is an impossibility. For this reason extensions to the quantum mechanical prescription have been given for constructing states given experimental observations of macroscopic systems.10) The most general of these prescriptions is that provided by the concept of a maximum-entropy estimate (hereafter referred to as MEE). This is based upon the existence of a unique measure (called the entropy) of the uncertainty represented by a probability distribution.* E. T. Jaynes11) has shown that a MEE provides a unique relationship between a density matrix and the experimental expectation

---

10) R. C. Tolman, op. cit.

*Shannon's (op. cit.) uniqueness theorem for the entropy is stated in Appendix A.

11) E. T. Jaynes, op. cit.

values of dynamic variables. He used this formalism to discuss the statistical mechanics of macroscopic systems. However the MEE method is not restricted to macroscopic systems. In this chapter we will apply it to a microscopic system.

The criterion of a MEE is to produce the least biased description of the manifold of states consistent with given expectation values of certain quantities. This means that no possible state of the system is ignored. A MEE assigns positive weight to every quantum state that is not absolutely excluded by the given information. For a microscopic system a complete measurement is possible. In this case the density matrix must reduce to one which describes a pure state. It is of interest to see how the MEE formalism handles this limiting case.

An Experiment on a Microscopic System

Consider an experimental observation of a non-relativistic free particle of spin one half. One possible complete set of commuting observables contains the momentum operator $\vec{p}$ and the z-component of spin angular momentum S_z. Another possible set is obtained by replacing the momentum operator by the position operator $\vec{x}$. However, in general an experimental observation of the system determines in some degree both position and momentum.

Suppose an experimental observation of the particle shows that

$$\langle \vec{X} \rangle = \langle \vec{p} \rangle = 0 \qquad (2\text{-}2a)$$

$$\langle x_i^2 \rangle = \Delta x_i^2 \qquad \langle p_i^2 \rangle = \Delta p_i^2 \qquad i = 1, 2, 3 \tag{2-2b}$$

$$\langle S_z \rangle = \hbar a \tag{2-2c}$$

Since $[x,p] = i\hbar$ there exists an uncertainty relation between the conjugate position and momentum coordinates of the particle.

$$\Delta x \, \Delta p \geq \frac{\hbar}{2} \tag{2-3}$$

A complete measurement can be characterized by the equalities

$$\Delta x \, \Delta p = \frac{\hbar}{2} \tag{2-4a}$$

$$a = \pm \frac{1}{2} \tag{2-4b}$$

For this limiting case the state of the system can be determined within the framework of quantum mechanics.[12]

$$\psi = c\, e^{-\frac{1}{4}\left(\frac{x_1^2}{\Delta x_1^2} + \frac{x_2^2}{\Delta x_2^2} + \frac{x_3^2}{\Delta x_3^2} \right)} \cdot \begin{cases} \alpha , & a = \frac{1}{2} \\ \beta , & a = -\frac{1}{2} \end{cases} \tag{2-5}$$

Where α and β are the spin-up and spin-down eigenstates of S_z. C is a normalization constant.

In the case of an inequality in (2-3) or in the case where $a \neq \pm\frac{1}{2}$, quantum mechanics does not provide a method of uniquely determining the wave function (or more generally the density matrix) corresponding to the experimental information Eq. (2-2). The problem of characterizing this initial state

---

[12] W. Pauli, Die Allgemeinen Prinzipien Der Wellenmechanik, (Reprinted from Handbuch der Physik, 2. Aufl., Band 24, 1. Tiel; J. W. Edwards Publisher, Ann Arbor, Michigan, 1950) p. 101.

is not merely an academic question. I. R. Senitzky[13)] in discussing noise limitations in microwave devices has shown that the predicted velocity distribution of an electron beam which passes through a microwave cavity depends upon the details of the wave packet representing the initially assumed electron state. The initial information concerning this state has the form of (2-2) with

$$\Delta x_i \, \Delta p_i \gg \frac{\hbar}{2}$$

The MEE formalism provides the only general method so far proposed for treating this type of problem.

The MEE prescription selects the density matrix ρ which conditionally maximizes the functional

$$S[\rho] = - Tr \rho \ln \rho$$

subject to the constraints imposed by the experimental observation (2-2)

$$Tr \rho x = Tr \rho p = 0$$

$$Tr \rho x^2 = \Delta x^2 \qquad Tr \rho p^2 = \Delta p^2$$

$$Tr \rho S_z = \hbar a$$

We shall consider the one dimensional case for simplicity. The results are easily generalized to the three dimensional case, since the space part of ρ factors into the direct product of three one-dimensional density matrices.

---

13) I. R. Senitzky, Phys. Rev. 95, 904 (1954).

Introducing Lagrange multipliers we seek to unconditionally maximize the functional

$$S[\rho] - \lambda_1 \mathrm{Tr}\,\rho x^2 - \lambda_2 \mathrm{Tr}\,\rho p^2 - \lambda_3 \mathrm{Tr}\,\rho S_z - \lambda_4 \mathrm{Tr}\,\rho x - \lambda_5 \mathrm{Tr}\,\rho p$$

The resulting MEE density matrix, normalized to unity is

$$\rho = \frac{\exp(-\lambda_1 x^2 - \lambda_2 p^2 - \lambda_3 S_z - \lambda_4 x - \lambda_5 p)}{Z(\lambda)}$$

with the partition function $Z(\lambda)$ given by

$$Z(\lambda) = \mathrm{Tr}\,\exp(-\lambda_1 x^2 - \lambda_2 p^2 - \lambda_3 S_z - \lambda_4 x - \lambda_5 p)$$

The λ parameters are determined by

$$\Delta x^2 = \langle x^2 \rangle = -\frac{\partial}{\partial \lambda_1} \ln Z \tag{2-6a}$$

$$\Delta p^2 = \langle p^2 \rangle = -\frac{\partial}{\partial \lambda_2} \ln Z \tag{2-6b}$$

$$\hbar a = \langle S_z \rangle = -\frac{\partial}{\partial \lambda_3} \ln Z \tag{2-6c}$$

$$0 = \langle x \rangle = -\frac{\partial}{\partial \lambda_4} \ln Z$$

$$0 = \langle p \rangle = -\frac{\partial}{\partial \lambda_5} \ln Z$$

Since the space and spin operators commute, Z may be factored into a product consisting of a spin part and a space part. The spin part of this product is

$$Z_{spin} = e^{-\lambda_3 \frac{\hbar}{2}} + e^{\lambda_3 \frac{\hbar}{2}} = 2 \cosh \frac{\hbar \lambda_3}{2} \tag{2-7}$$

In the space part we note by inspection that λ_4 and λ_5 must vanish to satisfy the zero expectation values of position and momentum. The trace is most easily evaluated in a representation where $\lambda_1 x^2 + \lambda_2 p^2$ is diagonal. The basis states for this representation are the harmonic oscillator wave functions.[14)]

$$(\lambda_1 x^2 + \lambda_2 p^2)|n\rangle = 2(\lambda_1\lambda_2)^{1/2}\hbar\,(n+1/2)\,|n\rangle$$

The space part of the trace becomes

$$Z_{space} = \sum_n e^{-2(\lambda_1\lambda_2)^{1/2}\hbar\,(n+1/2)} = \frac{1}{2\sinh \hbar(\lambda_1\lambda_2)^{1/2}} \quad (2\text{-}8)$$

The values of the λ parameters are found from (2-6).

$$\Delta X^2 = -\frac{\partial}{\partial\lambda_1}\ln Z = \frac{\hbar}{2}\left(\frac{\lambda_2}{\lambda_1}\right)^{1/2}\coth \hbar(\lambda_1\lambda_2)^{1/2}$$

$$\Delta p^2 = -\frac{\partial}{\partial\lambda_2}\ln Z = \frac{\hbar}{2}\left(\frac{\lambda_1}{\lambda_2}\right)^{1/2}\coth \hbar(\lambda_1\lambda_2)^{1/2}$$

$$\hbar a = -\frac{\partial}{\partial\lambda_3}\ln Z = -\frac{\hbar}{2}\tanh\frac{\hbar\lambda_3}{2}$$

Solving these equations leads to the relations,

$$\lambda_1 = \frac{1}{\hbar}\frac{\Delta p}{\Delta x}\coth^{-1}\frac{2}{\hbar}\Delta p\,\Delta x \quad (2\text{-}9a)$$

$$\lambda_2 = \frac{1}{\hbar}\frac{\Delta x}{\Delta p}\coth^{-1}\frac{2}{\hbar}\Delta p\,\Delta x \quad (2\text{-}9b)$$

$$\lambda_3 = \frac{2}{\hbar}\tanh^{-1}(-2a) \quad (2\text{-}9c)$$

Substituting these values into the partition function we find,

$$Z(\lambda) = \left[\frac{\left(\frac{2}{\hbar}\right)^2\Delta x^2\Delta p^2 - 1}{1-4a^2}\right]^{1/2} \quad (2\text{-}10)$$

The density matrix becomes

$$\rho = \left[\frac{1-4a^2}{\left(\frac{2}{\hbar}\right)^2\Delta x^2\Delta p^2 - 1}\right]^{1/2} e^{-\lambda_1 x^2 - \lambda_2 p^2 - \lambda_3 S_z}$$

---

14) L. I. Schiff, Quantum Mechanics, Second Edition (McGraw-Hill Book Company, Inc., New York, 1955) pp. 60, 62.

Now consider the probability which this density matrix assigns to the state $|n\pm\rangle$. This describes a spin-up (plus) or spin-down (minus) particle in the $n^{\underline{th}}$ harmonic oscillator space state.

$$\mathcal{P}(n,\pm) = \langle n\pm|\rho|n\pm\rangle = \left[\frac{1-4a^2}{(\frac{2}{\hbar})^2\Delta x^2\Delta p^2-1}\right]^{1/2} e^{-2(\lambda_1\lambda_2)^{1/4}\hbar(n+\frac{1}{2})\mp\lambda_3\frac{\hbar}{2}}$$

Using the relations (2-9) this becomes

$$\mathcal{P}(n,\pm) = \frac{1\pm 2a}{1+\frac{2}{\hbar}\Delta p\Delta x}\left(\frac{\frac{2}{\hbar}\Delta p\Delta x-1}{\frac{2}{\hbar}\Delta p\Delta x+1}\right)^n$$

This shows that n = 0 is always the most probable state. If $a > 0$ then the spin-up states have greater probability than the spin-down states and vice-versa.

Suppose the experiment provides a complete measurement of the system. That is Δx and Δp are measured as accurately as the Heisenberg uncertainty principle allows.

$$\Delta x\,\Delta p = \frac{\hbar}{2}$$

Furthermore

$$a = \tfrac{1}{2}$$

where we have arbitrarily assumed the system to have been found in the spin-up state. In this case

$$\mathcal{P}(n,+) = \delta_{n,0}$$

$$\mathcal{P}(n,-) = 0$$

Thus in this limiting case of a complete measurement, the MEE density matrix corresponds to the pure state

$$|0+\rangle = \frac{1}{(2\pi)^{1/4}(\Delta x)^{1/2}}\, e^{-\frac{x^2}{4\Delta x^2}} \cdot \alpha$$

This is just the result, Eq. (2-5), one obtains using the standard quantum mechanical formalism for constructing a state given a complete measurement.

It is interesting to observe that if $\Delta x \Delta p < \frac{1}{2}h$, the probabilities associated with the odd n states are negative. This is simply a manifestation of the Heisenberg uncertainty principle.

Finally, let us evaluate the entropy S associated with this density matrix.

$$S = -\operatorname{Tr}\rho \ln \rho$$

$$S = \ln z + \lambda_1 \Delta x^2 + \lambda_2 \Delta p^2 + \lambda_3 a$$

Using the relations (2-9) and (2-10) we find

$$S = \frac{b+1}{2}\ln\frac{b+1}{2} - \frac{b-1}{2}\ln\frac{b-1}{2} + \left(a-\tfrac{1}{2}\right)\ln(1-2a) - \left(a+\tfrac{1}{2}\right)\ln(1+2a) + \ln 2$$

where $b = \frac{2\Delta x \Delta p}{h}$. For b = 1 and $a = \pm\frac{1}{2}$ the entropy S is zero. This corresponds to a complete measurement. For an incomplete measurement in which

$$a = 0$$
$$\Delta x \Delta p \gg \frac{\hbar}{2}$$

the entropy becomes

$$S \doteq \ln b = \ln \frac{2\Delta x \Delta p}{\hbar} \equiv \ln W$$

So roughly the number of reasonably probable states for a spin one-half particle known to be in a region of phase space $\Delta x \Delta p$ is

$$W = \frac{2\Delta x \Delta p}{\hbar}$$

This is simply the statement that a quantum state corresponds to a phase volume of ħ per degree of freedom.

CHAPTER III

The High Probability Manifold Associated with a Maximum Entropy Density Matrix

In Appendix A three fundamental theorems of information theory are stated. In Chapter II the first of these theorems was used in the MEE construction of a density matrix. We now wish to investigate the physical interpretation of Theorems II and III. The information theory formulation of the second theorem is stated in terms of the probabilities associated with ergodic15) Markov sequences. In discussing the statistical mechanics of a macroscopic system we are interested in the probabilities associated with the global states of the system. For the limiting case of a macroscopic system composed of a large number of non-interacting particles, the eigenstates of any MEE density matrix based on expectation values of quantities of the form $F = \sum_i F_i$, where F_i operates only on the ith particle, factor into products of single particle states.16) For this idealized system, there exists a particularly simple, direct connection with the information theory theorems.

A global quantum state corresponds to an ergodic sequence, and the single particle states correspond to the symbols appearing in the sequence. Just as the sequences may be separated into two classes consisting of a high probability and a low probability group (Theorem II), the manifold of global

---

15) See Appendix A for the definition of ergodic.
16) As an example, consider a system of non-interacting spins located in an external magnetic field.

quantum states may be separated into a high probability manifold (HPM) and a low probability manifold (LPM). Furthermore, (Theorem III) a HPM can be defined such that all the states belonging to the HPM have nearly the same probability W^{-1} where W is the dimensionality of the HPM.

The importance of these theorems in the physical case is directly associated with the interpretation of the HPM belonging to a MEE density matrix. Since a MEE represents the least biased probability assignment consistent with the experimental conditions, it is surprising that it may actually specify a sharply defined manifold of high probability states. Furthermore, Theorem III shows that it assigns essentially the same probability to all the states in this manifold. This HPM must therefore represent the manifold of experimentally allowed quantum states for the system of interest. This identification forms a basis for discussing reversible and irreversible processes.17)

Shannon's proof18) of these theorems is based upon the assumed ergodic property of the symbol sequence. In the case of a system of interacting particles the global state wave function will not have this property. The interaction produces correlations between the single particle states which cannot be described by an ergodic Markov process. In order to clarify the effect of this on the existence of a HPM we will study a system of interacting spins. The spins are located on a rigid

---

17) See Chapter V.
18) C. E. Shannon, op. cit.

lattice and interact by means of a dipole-dipole interaction. The lattice is located in an external magnetic field H.

The general question of the existence of high probability manifolds for coupled spin systems has been discussed by R. Nelson[19]. Our treatment is not as general as Nelson's since a high temperature approximation is used. However, it is for a three dimensional system for which Nelson's results are not applicable. We reach the same conclusion as Nelson does for his one dimensional analysis. If the strength of the interaction falls off with distance rapidly enough, then a HPM exists. We have not found a general criterion for just how rapid this must be. It is sufficiently rapid if the correlations between spins produced by the interaction vanish for spins separated by more than some finite distance d. For the particular case considered here the criterion is that the effective local field H_L Eq. (3-14) be independent of the total number N of spins in the limit $N \to \infty$.

Definitions of a High Probability Manifold (HPM) and Separability

Consider a system of N interacting spins described by a Hamiltonian $\mathcal{H}$. An experimental observation gives $\langle \mathcal{H} \rangle = E$. The MEE density matrix is given by

$$\rho = \frac{e^{-\beta \mathcal{H}}}{Z(\beta)}$$

with

$$Z(\beta) = Tr\, e^{-\beta \mathcal{H}}$$

---

[19] R. Nelson, "The Statistical Basis of the Second Law," (Thesis, Stanford University Physics Department, 1961).

where β is determined by the condition

$$\mathrm{Tr}\,\rho\,\mathcal{H} = E$$

Each eigenstate of $\mathcal{H}$ represents a global state of the spin system. Let the degeneracy of the $n^{\underline{th}}$ energy eigenstate be denoted by g_n. The probability assigned an energy eigenstate $|n\rangle$ is,

$$P_m = \langle m|\rho|m\rangle = \frac{e^{-\beta E_m}}{Z(\beta)}$$

Order the states such that

$$g_1P_1 \geq g_2P_2 \geq \cdots \geq g_nP_n \geq g_{n+1}P_{n+1} \geq \cdots$$

Select in order from this series the degenerate eigenstates associated with g_1, g_2, ... Continue this selection until, for the first time, the sum of probabilities of the states selected is greater than or equal to λ $(0 < \lambda < 1)$.

$$\sum_{i=1}^{m(\lambda)} g_i P_i \geq \lambda \qquad (3\text{-}1)$$

The dimensionality of the manifold selected in this manner is

$$W(\lambda) = \sum_{i=1}^{m(\lambda)} g_i \qquad (3\text{-}2)$$

This manifold of states will be called the high probability manifold (HPM)[20]. The remaining states form the low probability manifold (LPM).

---

[20] This definition of the HPM excludes certain highly probable states (such as the ground state) which have small degeneracy factors. In this sense it differs from the definition of the high probability sequences given in Appendix A, Theorem II. It corresponds to the definition of HPM given by Theorem III, which is more appropriate physically.

Let $\bar{w}$ be defined by the relation

$$\lim_{N \to \infty} \frac{1}{N} \log W(\lambda) = \log \bar{w} \tag{3-3}$$

N represents the number of spins in the system. We shall say that the state space associated with a density matrix is separable (i.e. the HPM is sharply defined) if $\bar{w}$ is independent of λ $(0 < \lambda < 1)$.

The Method of Calculation

Define a density of states function N(E) by the relation

$$Z(\beta) = Tr\, e^{-\beta \mathcal{H}} = \int_{-\infty}^{\infty} dE\, N(E)\, e^{-\beta E} \tag{3-4}$$

For $\beta = 0$ we have the normalization condition

$$Tr\, 1 = \int_{-\infty}^{\infty} dE\, N(E) \tag{3-5}$$

For a system of non-interacting spins, located in an external field H, N(E) is a Gaussian centered about E = 0 and has an RMS width $\sqrt{kCH}$.[21] Where k is Boltzmann's constant, C is Curie's constant, and H is the external field.

We are interested in the case in which there exists a dipole-dipole interaction between the spins. Let the system contain N spins of angular momentum I located in a rigid lattice. The Hamiltonian consists of the Zeeman interaction with an external field H and the dipole-dipole interaction D.

$$\mathcal{H} = -HM + D \tag{3-6}$$

---

[21] See Appendix B.

where

$$M = \mu \sum_{i=1}^{N} I_{z_i}$$

$$D = \sum_{i>j} \sum_{\mu\nu} D_{\mu\nu}(ij)\, I_i^{\mu} I_j^{\nu}$$

Note that $\mathrm{Tr}\mathcal{H}$ = TrMD = 0. In Appendix B, N(E) is calculated for this system in the high temperature approximation. In this approximation we formally expand Z in powers of β. Then using Eq. (3-4) to define N(E) we have

$$\int_{-\infty}^{\infty} N(E) e^{-\beta E} dE = Z(\beta) \doteq \mathrm{Tr}\,1 \left(1 + \frac{\beta^2}{2} \frac{\mathrm{Tr}\mathcal{H}^2}{\mathrm{Tr}\,1}\right)$$

To this same order in β we find,

$$N(E) = \frac{N_0}{\pi^{1/2} \Delta^{1/2}} e^{-\frac{E^2}{\Delta}} \qquad (3\text{-}7)$$

where

$$N_0 = \mathrm{Tr}\,1 = (2I+1)^N$$

$$\Delta = \frac{2\,\mathrm{Tr}\mathcal{H}^2}{\mathrm{Tr}\,1} = 2kC\,(H^2 + H_L^2) \qquad (3\text{-}8a)$$

$$C = \frac{N\mu^2 I(I+1)}{3k} \qquad (3\text{-}8b)$$

$$H_L^2 = \frac{\mathrm{Tr}\,D^2}{\mathrm{Tr}\,M^2} \qquad (3\text{-}8c)$$

Compare (3-7) to the result for the non-interacting system of spins, Appendix B. We note that, in the approximation used, the effect of the interaction is simply to increase the width of the Gaussian by a factor proportional to $\frac{(H^2 + H_L^2)^{\frac{1}{2}}}{H}$.22)

---

22) From this vantage point it is not surprising that a HPM exists in the high temperature limit. The fact that only the average local field H_L enters the density of states function means in effect we are dealing with a system of "non-interacting" spins located in a magnetic field $(H^2 + H_L^2)^{\frac{1}{2}}$.

We now replace the summations in Eq's (3-1) and (3-2) by integrations over N(E). The only problem is to find the limits of integration. The energy E_1 corresponding to $g_1 P_1$ maximizes the function $N(E)e^{-\beta E}$.

$$\frac{d}{dE} N(E) e^{-\beta E} \Big|_{E=E_1} = 0$$

Using Eq. (3-7) one obtains

$$E_1 = -\frac{\beta\Delta}{2} \tag{3-9}$$

The sum in Eq's (3-1) and (3-2) extends to the set of degenerate states $n(\lambda)$. This corresponds in the continuous formulation to some energy $E_n(\lambda)$.

$$\frac{E^2}{\Delta} + \beta E = E_n(\lambda)$$

The roots of this equation are

$$E_{\pm} = -\frac{\beta\Delta}{2} \pm \sqrt{\left(\frac{\Delta\beta}{2}\right)^2 + E_n(\lambda)\Delta}$$

Define $\Gamma(\lambda)$ by

$$E_{\pm} = E_1 \pm \Delta^{1/2}\Gamma(\lambda) \tag{3-10}$$

Then Eq. (3-1) can be written as

$$\lambda = \frac{\int_{E_-}^{E_+} N(E) e^{-\beta E} dE}{\int_{-\infty}^{\infty} N(E) e^{-\beta E} dE} \tag{3-11}$$

where the limits of integration $E_{\pm}$ are given by (3-10). This determines $\Gamma(\lambda)$, Eq. (3-10), as a function of λ. Eq. (3-2) becomes

$$W(\lambda) = \int_{E_-}^{E_+} N(E)\, dE \tag{3-12}$$

The procedure is now to find $\gamma(\lambda)$ from (3-11), evaluate $W(\lambda)$ from (3-12), and then to investigate the limit

$$\lim_{N \to \infty} \frac{1}{N} \log W(\lambda)$$

Using the expression for N(E) (3-7) in (3-11) we find

$$\lambda = \frac{\int_{E_-}^{E_+} e^{-\frac{E^2}{\Delta} - \beta E}\, dE}{\int_{-\infty}^{\infty} e^{-\frac{E^2}{\Delta} - \beta E}\, dE}$$

Let $E = E_1 + \Delta^{\frac{1}{2}} x$ then

$$\lambda = \frac{1}{\pi^{1/2}} \int_{-\gamma(\lambda)}^{\gamma(\lambda)} e^{-x^2}\, dx = \operatorname{erf}(\gamma(\lambda))$$

This has the approximate solutions

$$\gamma(\lambda) = \begin{cases} \frac{\pi^{1/2}}{2} \lambda & \lambda \to 0 \\ \sqrt{-\ln(1-\lambda)} & \lambda \to 1 \end{cases}$$

The important point is that the N dependence in the limits $E_\pm$ is contained in the factor Δ, and $\gamma(\lambda)$ is finite for $\lambda \neq 1, 0$.

The dimensionality of the HPM is $W(\lambda)$. From (3-7) and (3-12) we have

$$W(\lambda) = \frac{N_0}{\pi^{1/2}\Delta^{1/2}} \int_{E_-}^{E_+} e^{-\frac{E^2}{\Delta}}\, dE$$

W(λ) can be bounded as follows[23)]

$$\frac{N_0}{\pi^{1/2}\Delta^{1/2}}\, e^{-\frac{E_-^2}{\Delta}}(E_+ - E_-) \leq W(\lambda) \leq \frac{N_0}{\pi^{1/2}\Delta^{1/2}}\, e^{-\frac{E_+^2}{\Delta}}(E_+ - E_-)$$

with

$$E_+ - E_- = 2\,\Delta^{1/2}\gamma(\lambda)$$

Therefore

$$\frac{N_0}{\pi^{1/2}}\, e^{-\frac{E_-^2}{\Delta}}\gamma(\lambda) \leq W(\lambda) \leq \frac{N_0}{\pi^{1/2}}\, e^{-\frac{E_+^2}{\Delta}}\gamma(\lambda)$$

Now assuming $\lambda \neq 1,0$ we obtain bounds on the $\lim\limits_{N\to\infty} \frac{1}{N} \log W(\lambda)$

$$\log(2I+1) - \lim_{N\to\infty} \frac{(E_1 - \Delta^{1/2}\gamma)^2}{N\Delta} \leq \lim_{N\to\infty} \frac{1}{N} \log W(\lambda) \leq \log(2I+1) - \lim_{N\to\infty} \frac{(E_1 + \Delta^{1/2}\gamma)^2}{N\Delta} \quad (3\text{-}13)$$

The λ-dependent term appearing in the lower bound of Eq. (3-13) is

$$\lim_{N\to\infty}\left(\frac{2E_1\gamma}{N\Delta^{1/2}} - \frac{\gamma}{N}\right) = \lim_{N\to\infty} \frac{2E_1\gamma}{N\Delta^{1/2}} = \lim_{N\to\infty} \beta\frac{\Delta^{1/2}\gamma}{2N}$$

The λ-dependent term in the upper bound of Eq. (3-13) has the opposite sign to this. We shall now examine the behavior of this term in the limit as N goes to infinity. Δ varies as

---

23) The point is that for a given λ we can find an N such that $E_- < E_+ < 0$. From the defining equation (3-10) for $E_\pm$ it follows that $E_- < E_+$. Now $E_+ = -\frac{\beta\Delta}{2} + \Delta^{\frac{1}{2}}\gamma(\lambda)$. For a given $\lambda \neq 1,0$, $\gamma(\lambda)$ is finite. Δ varies as N, therefore we may pick an N such that $E_+ < 0$. In particular this is true for $N > \frac{4\gamma^2(\lambda)}{\beta^2} \cdot \frac{3}{2\mu^2 I(I+1)(H^2 + H_L^2)}$.

$N(H^2+H_L^2)$. Therefore if H_L^2 is independent of N, the λ-dependent terms vanish as $N^{-\frac{1}{2}}$. However, if H_L^2 varies as N, the λ-dependent terms will not vanish in the limit of large N. Therefore the criterion for separability is that the square of the effective local field

$$H_L^2 = \frac{\mathrm{Tr}_\lambda D^2}{\mathrm{Tr}_\lambda M^2} \qquad (3\text{-}14)$$

be independent of N in the limit as N goes to infinity. This is the case for a dipole-dipole interaction between spins in a rigid lattice.

For the separable case the upper and lower bounds in Eq. (3-13) converge to the same limit. This gives

$$\log \bar{w} = \lim_{N\to\infty} \frac{1}{N} \log W(\lambda) = \log(2I+1) - \frac{\beta^2\mu^2 I(I+1)(H^2+H_L^2)}{6} \qquad (3\text{-}15)$$

The right-hand side is independent of λ and therefore the state space is separable. Furthermore, the entropy of the spin system $S = -\mathrm{Tr}\,\rho \ln \rho$, calculated to second order in β is

$$S = N\left[\log(2I+1) - \frac{\beta^2\mu^2 I(I+1)(H^2+H_L^2)}{6}\right]$$

Therefore

$$\log \bar{w} = \lim_{N\to\infty} \frac{1}{N} \log W(\lambda) = \frac{S}{N} \qquad (3\text{-}16)$$

and $\log \bar{w}$ is the entropy per spin. This corresponds to Theorem II in Appendix A.

We shall next investigate the probability associated with a state belonging to the HPM. The energy of a HPM state must lie between E_+ and E_-, Eq. (3-10). Therefore the probability associated with a HPM state is bounded by the Boltzmann factors,[24]

$$\frac{e^{-\beta(E_1-\Delta^{1/2}\gamma)}}{(2I+1)^N e^{\frac{\beta^2\Delta}{4}}} \le P \le \frac{e^{-\beta(E_1+\Delta^{1/2}\gamma)}}{(2I+1)^N e^{\frac{\beta^2\Delta}{4}}}$$

Next we note that

$$\log\left[\frac{e^{-\beta E_1}}{(2I+1)^N e^{\frac{\beta^2\Delta}{4}}}\right]^{-1} = S$$

Furthermore, E_1 varies as N and $\Delta^{\frac{1}{2}}$ varies as $N^{\frac{1}{2}}$. Therefore, given $\delta > 0$, we can find an N' such that for $N > N'$

$$\left|\frac{\log P^{-1} - S}{N}\right| < \delta$$

In this limit we have

$$P = \overline{W}^{-N}$$

This result corresponds to Theorem III in Appendix A.

We may summarize these results by observing that if the state space associated with a density matrix is separable, then it is possible to treat the states as though there were just $\overline{W}^N$ of them each having probability $\overline{W}^{-N}$. The entropy S determines the size of $\overline{W}$.

$$\overline{W} = e^{\frac{S}{N}}$$

For a MEE density matrix, corresponding to no experimental information concerning the state of the system, S = N log (2I + 1) and $\overline{W} = 2I + 1$. For a MEE density matrix, corresponding to some experimental observation having bearing on

---

24) See footnote 23.

the state of the system, $S < N \log(2I + 1)$ and $\overline{W} < (2I + 1)$. In this case we note that in the limit of large N the HPM of states corresponds to a negligibly small fraction of the possible global states of the system. It is just this fraction of the states which we must deal with in analyzing experimental results.

CHAPTER IV

MEE for Stationary Problems

The problem of theoretical physics is the prediction of experimentally reproducible results. Reproducibility means that for almost all of the quantum states allowed by the experimental control, the observed results are similar. This does not exclude fluctuations; in fact, it makes them meaningful. This macroscopic uniformity[25)] means that reproducible experimental results may be calculated by averaging over the manifold of experimentally allowed states. Fluctuations may of course be calculated in this same manner. Therefore, the problem is simply to construct this manifold corresponding to a given experimental set up. The MEE formalism should provide a method for doing this.

As discussed in Appendix A it does provide a unique method for constructing a probability distribution. This is unique in the sense that it is the least biased distribution consistent with the given experimental information. Furthermore, this density matrix characterizes a separation of the states of the system into two manifolds. One manifold (the HPM) contains states, the sum of whose probabilities approaches

---

25) J. Von Neumann, Z. Physik 57, 30 (1929). P. Bocchieri and A. Loinger, Phys. Rev. 114, 948 (1958) repeat Von Neumann's argument after making a unitary transformation. They show that for the "overwhelming majority" of the macroscopically allowed initial states of the system the expectation values of the dynamic variables of interest are the same. This is what we refer to as macroscopic uniformity.
See also E. T. Jaynes, Phys. Rev. 108, 171 (1957).

arbitrarily close to unity. The dimensionality W of this manifold is

$$\log W = S$$

where S is the entropy of the distribution. The probability associated with any state belonging to the HPM is essentially $\frac{1}{W}$. Now consider what these facts imply. The least biased probability assignment consistent with the experimental arrangement characterizes a HPM of uniformly probable states. This must therefore correspond to the experimentally allowed manifold of states. *Thus we have a principle for constructing this manifold.*

In this chapter we shall begin an investigation of the manifolds associated with stationary experiments. In a stationary experiment the values of the observed quantities do not depend upon the time. A special class of stationary experiments are those of thermodynamic equilibrium. For this case Jaynes[26] has shown that the MEE formalism reduces to the standard computational rules of equilibrium statistical mechanics. Specifically, if the information characterizing the experimentally allowed states is the average energy, the MEE density matrix has the canonical form $e^{\beta(F-\mathcal{H})}$.

Now consider a more general type of stationary experiment. For example, let us suppose that steady, external constraints (concentration gradients, temperature gradients) prevent a system from reaching thermal equilibrium. At present there

---

26) E. T. Jaynes, op. cit.

does not exist a computational formalism equivalent to equilibrium statistical mechanics for this type of stationary process. We believe MEE provides the basis for such a formalism. The problem is two-fold. We must correctly enumerate the relevant experimental information; and secondly, we must calculate the MEE density matrix consistent with this information.

In order to investigate the possible methods of characterizing stationary experiments, we shall first consider a thermal equilibrium problem. The advantage of this is that the correct results are known from equilibrium statistical mechanics. For example, the equation of state for a spin system establishes an equilibrium relationship between H, M and T. The statement that a unique equation of state exists means, experimentally, that given any two of the quantities H, M, T, the third is determined. Thus, in a consistent theory, the same equation of state must emerge whether the MEE density matrix is based on knowledge of (T,H), (H,M), or (M,T). Conventional statistical mechanics considers only the first of these cases. However, in the latter cases it is necessary to explicitly state (in some form) that the physical situation we wish to describe is stationary in time. Analysis of this problem leads to a method of imposing the stationary constraint.

The Equation of State for a Spin System

One of the properties characterizing a thermodynamic system is an experimentally reproducible equation of state.

Consider a system of spins located on a rigid lattice and interacting by means of a dipole-dipole interaction. For this system we expect an equation of state establishing the relationship between T, H and M.

$$f(T,H,M) = 0$$

Usually one considers T and H as given. The density matrix describing this situation is

$$\rho = e^{-\lambda_0 - \beta \mathcal{H}} \tag{4-1}$$

$$\beta = \frac{1}{kT} \qquad \mathcal{H} = -HM + D$$

$$\lambda_0 = \ln Z \qquad Z = \mathrm{Tr}\, e^{-\beta \mathcal{H}}$$

where D represents the dipole-dipole interaction. The equation of state is obtained by calculating the expectation value of the magnetization.

$$\langle M \rangle = \mathrm{Tr}\, \rho M$$

In lowest order, the high temperature expansion of the equation of state is Curie's Law.

$$\langle M \rangle = \mathrm{Tr}\, \rho M \doteq \frac{CH}{T} \tag{4-2}$$

$$C = \frac{1}{k} \frac{\mathrm{Tr}\, M^2}{\mathrm{Tr}\, 1} = \frac{N \mu^2 I(I+1)}{3k}$$

We assume this equation of state is a reproducible experimental result. Therefore, this value of $\langle M \rangle$ must be not only the average obtained from the density matrix (4-1) but the value characteristic of each of the great majority of the eigenstates in the HPM corresponding to (4-1). This HPM corresponds physically to the manifold of experimentally allowed states. Now consider the case in which the values of the external field H and the magnetization M are given. This information, coupled with a correct statement of the stationary nature of the equilibrium situation, must characterize this same experimentally allowed manifold of states.

To verify this we shall construct the maximum entropy density matrix consistent with the $\langle M \rangle$, H information. A determination of the temperature associated with this density matrix yields an equation of state. Under approximations equivalent to the high temperature expansion this equation of state should be just the Curie Law, Eq. (4-2). The density matrix will, in general, not be canonical. Therefore, in order to proceed we need an operational definition of the temperature associated with a manifold of states. In the case of a canonical density matrix the temperature is defined in terms of a parameter which multiplies $\mathcal{H}$. Specifically one has

$$\frac{1}{T} = \frac{d\,S(\langle\mathcal{H}\rangle)}{d\langle\mathcal{H}\rangle} \qquad (4\text{-}3)$$

A natural generalization of this is provided by Gibbs' Second Analogy definition of temperature.27)

---

27) J. Willard Gibbs, Elementary Principles in Statistical Mechanics, (Dover Publications, Inc., New York, N. Y., 1960; originally published by the Yale University Press, New Haven, Conn. 1902) Chapter XIV.

$$\frac{1}{T} = \frac{d}{dE} k \log N(E) \Big|_{E = \langle \mathcal{H} \rangle} \qquad (4\text{-}4)$$

N(E) represents the energy density of states function and k is Boltzmann's constant. $\langle \mathcal{H} \rangle$ is the expectation value of the energy. This expectation value may be defined over any manifold of states. In particular, Eq. (4-4) can be used to operationally define the temperature of a single quantum state. Therefore, by using Eq. (4-4) we may say that temperature is a property of a single, global, quantum state. When temperature is a useful parameter in correlating experimental observation, Eq. (4-4) will agree with experimental results. For a canonical density matrix which is separable (i.e. a sharply defined HPM exists), the temperature defined by Eq. (4-4) agrees with that of Eq. (4-3) to order $N^{-\frac{1}{2}}$. Where N is proportional to the number of degrees of freedom of the system.

To begin with, let us consider the MEE density matrix consistent with the expectation value of the magnetization $\langle M \rangle$ and the value of the external field H. It has the form

$$\begin{aligned} \rho &= e^{-\lambda_0 - \lambda_1 M} \\ Z &= \mathrm{Tr}\, e^{-\lambda_1 M} \\ \lambda_0 &= \ln Z \end{aligned} \qquad (4\text{-}5)$$

and λ_1 is determined by the requirement

$$\langle M \rangle = -\frac{\partial}{\partial \lambda_1} \ln Z$$

For a stationary situation, $\langle M \rangle$ should be independent of time. Using the density matrix (4-5) we find

$$\frac{d}{dt}\langle M\rangle = \frac{i}{\hbar}\, Tr\, [\rho, \mathcal{H}]\, M = 0$$

$$\frac{d^2}{dt^2}\langle M\rangle = -\frac{1}{\hbar^2}\, Tr\, [[\rho, \mathcal{H}], \mathcal{H}]\, M \neq 0$$

Further, higher order time derivatives will also not vanish. Therefore, the density matrix (4-5) does not describe a stationary experiment.[28)]

Therefore, we must find a method of explicitly stating the information that the experimental observables are stationary. One possibility is to restrict the admissible density matrices to the class which commute with the Hamiltonian of the system.

$$[\mathcal{H}, \rho] = 0 \qquad (4\text{-}6)$$

We shall call Eq. (4-6) the "strong stationary condition." If it is satisfied then the expectation value of the magnetization will be stationary. Moreover, the expectation values of any dynamic variable will be stationary.

The MEE density matrix will be constructed by maximizing

$$S[\rho] = -\, Tr\, \rho \ln \rho$$

---

28) Had we continued the calculation (evaluating $\langle \mathcal{H} \rangle$, and then T using Eq. (4-4)) we would have found, keeping terms of order H_L^2

$$\frac{1}{T} = \frac{\langle M\rangle}{CH}\left(1 - \frac{H_L^2}{H^2}\right) \qquad H_L^2 = \frac{Tr\, D^2}{Tr\, M^2}$$

This is not the correct equilibrium equation of state.

subject to the constraints

$$\mathrm{Tr}\,\rho = 1$$
$$\mathrm{Tr}\,\rho M = \langle M \rangle$$

and the "strong stationary condition"

$$[\rho, \mathcal{H}] = 0 \qquad \mathcal{H} = -HM + D$$

where $\langle M \rangle$ and H are known. The analysis is most conveniently carried out in a representation in which $\mathcal{H}$ is diagonal.

$$\mathcal{H}|m\rangle = E_m |m\rangle$$

ρ will be stationary if it has the form

$$\langle m|\rho|m\rangle = \langle m|\rho|m\rangle\, \delta_{mm}$$

Assuming this form for ρ, we maximize the expression

$$S - \lambda \langle M \rangle$$

The result is,[29]

$$\langle m|\rho|m\rangle = e^{-\lambda_0 - \lambda_1 \langle m|M|m\rangle} \tag{4-7}$$

$$Z = \sum_m e^{-\lambda_1 \langle m|M|m\rangle}$$

$$\lambda_0 = \ln Z \qquad \langle M \rangle = -\frac{\partial}{\partial \lambda_1} \ln Z$$

From ρ we may calculate $\langle \mathcal{H} \rangle$, then using Eq. (4-4) we calculate the temperature associated with ρ. This determines the equation of state for the spin system.

---

29) A detailed analysis of this variational problem is given later in this chapter where the general stationary problem is discussed.

We shall use a formal expansion of the exponential. When the calculation is completed it will be clear that this is an expansion in the ratio of an energy per spin to kT.30) Throughout the calculation the states $|n\rangle$ will be the eigenstates of the total Hamiltonian $\mathcal{H}$. Using the density matrix given in Eq. (4-7) the expectation value for the magnetization is

$$\langle M\rangle = \frac{\sum_m \langle m|M|m\rangle e^{-\lambda_1 \langle m|M|m\rangle}}{\sum_m e^{-\lambda_1 \langle m|M|m\rangle}}$$

Expanding the exponential and keeping only the lowest non-vanishing contribution to $\langle M\rangle$ we obtain

$$\langle M\rangle \doteq -\lambda_1 \frac{\sum_m \langle m|M|m\rangle^2}{Tr\, 1}$$

The zero order term varies as TrM which vanishes. This equation determines the value of the parameter λ_1, in terms of the expectation value of the magnetization.

$$\lambda_1 = -\langle M\rangle \frac{Tr\, 1}{\sum_m \langle m|M|m\rangle^2} \qquad (4\text{-}8)$$

Again using the density matrix (4-7), the expectation value for the total energy of the spin system is

$$\langle \mathcal{H}\rangle = \frac{\sum_m \langle m|\mathcal{H}|m\rangle e^{-\lambda_1 \langle m|\mathcal{H}|m\rangle}}{\sum_m e^{-\lambda_1 \langle m|M|m\rangle}}$$

---

30) The expansion parameter in the high temperature approximation is the ratio of the energy per spin to kT. See footnote p. 1092, J. H. Van Vleck, Supp. Vol. VI, Serie X, Del Nuovo Cimento (1957). For the example which we consider $uH/kT \sim 10^{-3}$ for $H \sim 10^4$ gauss and $T \sim 1$°K.

Expanding the exponential we obtain in lowest order

$$\langle \mathcal{H} \rangle \doteq \frac{-\lambda_1 \sum_n \langle n|\mathcal{H}|n\rangle \langle n|M|n\rangle}{Tr\,1}$$

The zero order term varies as Tr$\mathcal{H}$ which vanishes. Since is diagonal in the $|n\rangle$ representation we may write the last equation for $\langle \mathcal{H} \rangle$ in terms of a trace.

$$\langle \mathcal{H} \rangle \doteq -\lambda_1 \frac{Tr\,\mathcal{H} M}{Tr\,1}$$

Since TrMD vanishes this last expression becomes

$$\langle \mathcal{H} \rangle \doteq \lambda_1 H \frac{Tr\,M^2}{Tr\,1} \tag{4-9}$$

Finally, substituting the expression for λ_1 given in (4-8) we obtain

$$\langle \mathcal{H} \rangle \doteq -H \cdot \langle M \rangle \frac{Tr\,M^2}{\sum_n \langle n|M|n\rangle^2} \tag{4-10}$$

In the limit D → 0 we get the expected result,

$$\langle \mathcal{H} \rangle = -H \cdot \langle M \rangle$$

In order to evaluate $\langle \mathcal{H} \rangle$ we must calculate the sum

$$\sum_n \langle n|M|n\rangle^2 \tag{4-11}$$

We shall seek to cast this sum (4-11) into the form of a trace. The advantage of this is the well known fact that the value of a trace is representation independent. In particular we may use the representation in which M is diagonal. Let these states be represented by $|m_\circ\rangle$,

$$M|m_\circ\rangle = |m_\circ\rangle M_{m_\circ}$$

The exact energy eigenstate $|n\rangle$ may be expanded in terms of the $|n_0\rangle$ basis states,

$$|n\rangle \sim |n_0\rangle + \sum_{n_0'}{}' |n_0'\rangle\langle n_0'| \frac{1}{E_{n_0}-E_{n_0'}} D\,|n_0\rangle + \cdots$$

Only the lowest order D correction to $|n_0\rangle$ will be kept. This is consistent with the fact that in calculating $\langle M\rangle$ and $\langle \mathcal{H}\rangle$ only the lowest order contributions were retained. Let the $|n_0\rangle$ eigenfunctions be normalized to unity. Then

$$\langle n|\,M\,|n\rangle \doteq \langle n_0|M|n_0\rangle + \sum_{n_0'}{}' \frac{|\langle n_0|D|n_0'\rangle|^2}{(E_{n_0}-E_{n_0'})^2}\Big(\langle n_0'|M|n_0'\rangle - \langle n_0|M|n_0\rangle\Big)$$

Now compute the square of the matrix element $\langle n|M|n\rangle$, keeping terms of order D^2 in the interaction.

$$\langle n|M|n\rangle^2 \doteq \langle n_0|M|n_0\rangle^2$$

$$+\,2\sum_{n_0'}{}' \frac{|\langle n_0|D|n_0'\rangle|^2}{(E_{n_0}-E_{n_0'})^2}\Big(\langle n_0'|M|n_0'\rangle - \langle n_0|M|n_0\rangle\Big)\langle n_0|M|n_0\rangle$$

The sum (4-11) becomes, to order D^2 in the interaction,

$$\sum_{n}\langle n|M|n\rangle^2 \doteq \sum_{n_0}\langle n_0|M|n_0\rangle^2 + 2\sum_{n_0 n_0'}{}' \frac{|\langle n_0|D|n_0'\rangle|^2}{(E_{n_0}-E_{n_0'})^2}\Big(\langle n_0'|M|n_0'\rangle\langle n_0|M|n_0\rangle - \langle n_0|M|n_0\rangle^2\Big)$$

The last term may be written as,

$$\sum_{n_0 n_0'}{}' \frac{\langle n_0|[D,M]|n_0'\rangle\langle n_0'|[D,M]|n_0\rangle}{(E_{n_0}-E_{n_0'})^2} = -\frac{1}{H^2}\,\mathrm{Tr}\,D^2$$

The first term is simply $\mathrm{Tr}M^2$. Therefore the sum (4-11) can be cast in the form,

$$\sum_{m} \langle m|M|m\rangle^2 \doteq \mathrm{Tr}\, M^2 - \frac{1}{H^2}\, \mathrm{Tr}\, D^2 \qquad (4\text{-}12)$$

Using this in (4-10) one obtains

$$\langle \mathcal{H} \rangle \doteq -H \cdot \langle M \rangle \, \frac{\mathrm{Tr}\, M^2}{\mathrm{Tr}\, M^2 - \frac{1}{H^2}\, \mathrm{Tr}\, D^2}$$

Define a local field

$$H_L^2 = \frac{\mathrm{Tr}\, D^2}{\mathrm{Tr}\, M^2}$$

then,

$$\langle \mathcal{H} \rangle \doteq \frac{-H \cdot \langle M \rangle}{1 - \left(\frac{H_L}{H}\right)^2} \qquad (4\text{-}13)$$

The high temperature approximation to the energy density of states N(E) has been given, Eq. (3-7). Using the definition of T Eq. (4-4) we find

$$\frac{1}{T} = \frac{d}{dE}\, k \log N(E) \Big|_{E = \langle \mathcal{H} \rangle} = \frac{-2 \langle \mathcal{H} \rangle}{\Delta}$$

Substituting the expectation value of $\mathcal{H}$, Eq. (4-13), this becomes,

$$\frac{1}{T} = \frac{H \cdot \langle M \rangle}{c\left(1 - \left(\frac{H_L}{H}\right)^2\right)\left(H^2 + H_L^2\right)} = \frac{\langle M \rangle}{cH}\left[1 - \left(\frac{H_L}{H}\right)^4\right]^{-1}$$

To second order in D^2 ($H_L^2 = \mathrm{Tr}D^2/\mathrm{Tr}M^2$) we obtain

$$\frac{1}{T} = \frac{\langle M \rangle}{cH}$$

Therefore the equation of state giving T in terms of $\langle \underline{M} \rangle$ and H is Curie's Law

$$T = \frac{C\,H}{\langle M \rangle}$$

This is exactly the equation of state obtained in the high temperature approximation from the density matrix characterized by the initial information T,H.

Since the "strong stationary condition" Eq. (4-6) reproduces the Curie Law for this equilibrium problem, we shall use it to investigate non-equilibrium problems. A general formalism using Eq. (4-6) to specify the stationary character of the physical problem is developed. The concept of the diagonal part of an operator allows the density matrix to be written in a representation independent form (as opposed to Eq. (4-7)). An explicit representation for the diagonal part of an operator allows us to calculate with this formalism.

A General Stationary Formalism Using the Strong Stationary Condition

Although the previous example concerned a thermal equilibrium problem, the analysis was based upon only the stationary character of the observations. This suggests that we extend this same method to stationary, non-equilibrium problems.

Suppose we are given the expectation values of the operators F_1, $F_2 \ldots F_n$. The density matrix which represents the most unbiased description of the state of the system consistent with these conditions is

$$\rho = e^{-\lambda_0 - \sum_i \lambda_i F_i} \tag{4-14}$$

where λ_0 and λ_i are determined by

$$Z = \mathrm{Tr}\, e^{-\sum_i \lambda_i F_i}$$

$$\lambda_0 = \ln Z$$

$$\langle F_i \rangle = -\frac{\partial}{\partial \lambda_i} \ln Z$$

If any of the F_i fail to commute with the Hamiltonian, the density matrix (4-14) is not stationary. Its time dependence is determined by the Liouville equation

$$i\dot{\rho} = [\mathcal{H}, \rho] \qquad (\hbar = 1 \text{ in this section})$$

Suppose besides being given the above expectation values of F_i we are interested in describing a stationary experiment. A weak condition characterizing a stationary experiment is given by

$$\frac{d}{dt} \mathrm{Tr}\, \rho\, F_i(t) = 0 \tag{4-15}$$

with

$$F_i(t) = e^{i\mathcal{H}t} F_i\, e^{-i\mathcal{H}t}$$

Sufficient conditions for (4-15) are

$$[\mathcal{H}, F_i] = 0 \qquad \text{or} \qquad [\mathcal{H}, \rho] = 0$$

We shall be interested in cases where $[\mathcal{H}, F_i] = 0$. As before, we will assume the "strong stationary condition."

$$[\mathcal{H}, \rho] = 0 \tag{4-16}$$

We now calculate ρ by maximizing the entropy $S = -\mathrm{Tr}\rho\ln\rho$ over the manifold of density matrices which commute with $\mathcal{H}$, and which satisfy the constraints,

$$\mathrm{Tr}\,\rho F_i = \langle F_i \rangle \tag{4-17}$$

Proceeding in the standard manner let

$$J[\rho] = \mathrm{Tr}\,\rho\ln\rho + \sum_i \lambda_i \mathrm{Tr}\,\rho F_i$$

Let $\rho = \bar{\rho} + \epsilon\eta$, $[\bar{\rho}, \mathcal{H}] = 0$, where $\bar{\rho}$ is the density matrix which makes J stationary. η is an arbitrary variation subject to the constraint $[\eta, \mathcal{H}] = 0$. Then taking the first variation,

$$\frac{d}{d\epsilon} J[\bar{\rho} + \epsilon\eta] \Big|_{\epsilon = 0} = 0$$

we find*

$$\mathrm{Tr}\,\eta\,[\ln\bar{\rho} + \sum_i \lambda_i F_i] = 0 \tag{4-18}$$

If η were completely arbitrary, one would require the operator identity $\ln\bar{\rho} + \sum_i \lambda_i F_i = 0$. However, we have the constraint $[\eta, \mathcal{H}] = 0$. Let us write this constraint condition in a representation which diagonalizes $\mathcal{H}$.

$$\mathcal{H}|E\nu\rangle = |E\nu\rangle E \tag{4-19}$$

Where ν represents the additional quantum numbers associated with the state.

$$\langle E\nu|[\mathcal{H}, \eta]|E'\nu'\rangle = 0$$

$$(E - E')\langle E\nu|\eta|E'\nu'\rangle = 0 \tag{4-20}$$

---

*An additive term +1 has been omitted inside the bracket, Eq. (4-18), since ρ will be normalized at the end of the calculation.

When $E \neq E'$, the matrix element $\langle E\nu|\eta|E'\nu'\rangle$ vanishes. However, within any degenerate manifold of $\mathcal{H}$ the matrix elements of η may be arbitrary. The condition on $\bar{\rho}$ given by equation (4-18) is therefore,

$$\langle E\nu|\ln\bar{\rho} + \sum_i \lambda_i F_i |E\nu'\rangle = 0 \tag{4-21}$$

Let us define the "diagonal part" of F_i by the relation

$$\langle E\nu|F_{d_i}|E'\nu'\rangle = \delta_{EE'}\langle E\nu|F|E\nu'\rangle \tag{4-22}$$

Since $\bar{\rho}$ is diagonal in E, we may write Eq. (4-21) in the form

$$\langle E\nu|\ln\bar{\rho} + \sum_i \lambda_i F_{d_i}|E'\nu'\rangle = 0 \tag{4-23}$$

Since the states $|E\nu\rangle$ form a complete set, we have the operator relation

$$\ln\bar{\rho} + \sum_i \lambda_i F_{d_i} = 0 \tag{4-24}$$

The steady state MEE density matrix normalized to unity is therefore

$$\bar{\rho} = \frac{e^{-\sum_i \lambda_i F_{d_i}}}{Z(\lambda_i)} \qquad Z(\lambda_i) = \mathrm{Tr}\, e^{-\sum_i \lambda_i F_{d_i}} \tag{4-25}$$

We shall now proceed to construct an explicit representation for the diagonal operator F_d. Let

$$F_d = F - i\,[\mathcal{H}, A] \tag{4-26}$$

then we have satisfied, for well behaved A, the condition,

$$\langle E\nu|F_d|E\nu'\rangle = \langle E\nu|F|E\nu'\rangle$$

In order to remove the non-diagonal part of F we require

$$\langle E\nu|F|E'\nu'\rangle = i\,\langle E\nu|[\mathcal{H}, A]|E'\nu'\rangle \qquad E \neq E'$$

Therefore let A have the off diagonal matrix elements

$$\langle E\nu | A | E'\nu' \rangle = \frac{\langle E\nu | F | E'\nu' \rangle}{i(E-E')} \qquad E \neq E' \tag{4-27}$$

A formal representation for A is

$$A = \lim_{\epsilon \to 0} \int_{-\infty}^{0} dt\, e^{\epsilon t} e^{i\mathcal{H}t} F e^{-i\mathcal{H}t} \tag{4-28}$$

Taking matrix elements we have

$$A = \lim_{\epsilon \to 0} \frac{\langle E\nu | F | E'\nu' \rangle}{i(E-E') + \epsilon} = \frac{\langle E\nu | F | E'\nu' \rangle}{i(E-E')}$$

Therefore the diagonal part of an operator can be represented by

$$F_d = F - \int_{-\infty}^{0} dt\, e^{\epsilon t} \dot{F}(t) \qquad F(t) = e^{i\mathcal{H}t} F e^{-i\mathcal{H}t} \tag{4-29}$$

Where, in order to ease the notation, we have omitted "$\lim_{\epsilon \to 0}$". This limiting process is understood to occur at the end of the calculation. Actually it is sufficient to let ϵ remain finite if it is small compared to the energy level spacing. In the limit of continuous energy levels ϵ must go to zero.

In order to gain a more physical picture of the "diagonal part" of an operator, consider Eq. (4-29). Let us treat F(t) as a c-number function having a d.c. part F_0. Carrying out the integration we obtain in the limit ϵ goes to zero

$$F_d = F_0 = \lim_{\epsilon \to 0} \epsilon \int_{-\infty}^{0} F(t) e^{\epsilon t} dt$$

Therefore the "diagonal part" of an operator is that part which remains constant under a unitary transformation generated by the Hamiltonian $\mathcal{H}$. It is just this time independent part which one would expect to play a role in stationary processes.

As an example of this formalism we consider the problem of steady-state diffusion of particles through a lattice. For small concentration gradients we obtain an expression for the diffusion constant. This is identical to that obtained by R. Kubo et. al. using a different method.31) Higher order approximations of the stationary MEE formalism provide a prescription for calculating non-linear effects in steady-state problems. Kubo's formalism is not applicable to this problem. He assumes that the constraints preventing the system from reaching thermodynamic equilibrium are removed at t = 0. His results are based upon the time development of the density matrix as it evolves toward equilibrium.32)

<u>The Problem of Steady-State Diffusion</u>

Consider the problem of steady-state diffusion of particles through a lattice. We assume that the observed quantities characterizing this problem are the temperature of the system T and the expectation value of the concentration of particles $\langle n(x) \rangle$. From a conceptual point of view, the second quanti-

---

31) R. Kubo et. al. op. cit. This method is based upon Onsager's assumption that the average regression of spontaneous fluctuations follow the macroscopic laws of irreversible processes. For a calculation more closely related to ours, see S. Nakajima, Prog. Theor. Phys. (Japan) <u>20</u>, 948 (1958).

32) For higher order approximation in Kubo's formalism see W. Bernard and H. B. Callen, Rev. Mod. Phys. <u>31</u>, 1017, (1959).

zation formalism is particularly appropriate to this problem since any attempt to describe a state of steady flow in terms of configuration space would involve particle coordinates appearing and disappearing from the region of interest. In addition, the second quantization formalism automatically takes account of the Bose or Fermi statistics of identical particles, whereas in configuration space treatments this is accomplished only be imposing additional arbitrary constraints on the problem. Let the Hamiltonian for the system be

$$\mathcal{H} = \int d^3x\, \psi^+(x) \left(-\frac{\hbar^2}{2m}\nabla^2 + V(x)\right)\psi(x) \tag{4-30}$$

The density of particles operator is

$$n(x) = \psi^+(x)\,\psi(x) \tag{4-31}$$

The continuity equation is

$$\dot{n} + \vec{\nabla}\cdot\vec{j} = 0 \tag{4-32}$$

with the current operator defined by the relation

$$\vec{j} = \mathrm{Re}\ \psi^+ \frac{\hbar}{im}\vec{\nabla}\psi \tag{4-33}$$

Given T and $\langle n(x)\rangle$ the steady-state MEE density matrix is, Eq. (4-26)

$$\rho = \frac{e^{-\beta\mathcal{H} + \int d^3x\,\mu(x)\,n_d(x)}}{Z_\beta[\mu(x)]} \tag{4-34}$$

where $n_d(x)$ is the diagonal part of the operator n(x) and

$$Z_\beta[\mu(x)] = \mathrm{Tr}\, e^{-\beta\mathcal{H} + \int d^3x\,\mu(x)\,n_d(x)} \tag{4-35}$$

$Z_\beta[\mu(x)]$ is a function of β and a functional of $\mu(x)$.33) β equals $(kT)^{-1}$ and the chemical potential $\mu(x)$ is determined by the condition

$$\langle n(x)\rangle = \frac{\delta \ln Z_\beta[\mu(x)]}{\delta \mu(x)} \qquad (4\text{-}36)$$

where $\frac{\delta}{\delta\mu(x)}$ represents a variational derivative.

The expectation value of the current density $\vec{j}(x)$ is

$$\langle j(x)\rangle = \mathrm{Tr}\, \rho\, \vec{j}(x) \qquad (4\text{-}37)$$

Assuming a small density gradient we expand the density matrix keeping only the lowest order term

$$\langle \vec{j}(x)\rangle \doteq \mathrm{Tr}\left(\rho_0 \left[1 + \int_0^\beta \frac{ds}{\beta} \int d^3x'\, \mu(x')\, n_d(x', -is)\right] \vec{j}(x)\right) \qquad (4\text{-}38)$$

where

$$n_d(x', -is) = e^{\mathcal{H}s}\, n_d(x')\, e^{-\mathcal{H}s}$$

$$\rho_0 = \frac{e^{-\beta\mathcal{H}}}{Z(\beta)} \qquad Z(\beta) = \mathrm{Tr}\, e^{-\beta\mathcal{H}}$$

---

33) In the case where the local temperature T(x) is a function of position

$$\rho = \frac{\exp\left(-\int d^3x \left[\beta(x)\, \psi^\dagger(x)\left(-\frac{\nabla^2}{2m} + V\right)\psi(x) - \mu(x)\, n_d(x)\right]\right)}{Z[\beta(x), \mu(x)]}$$

and Z is a functional of $\beta(x)$ and $\mu(x)$.

$$\beta(x) = \frac{1}{kT(x)} \qquad \langle n(x)\rangle = \frac{\delta \ln Z}{\delta \mu(x)}$$

For a discussion of the partition functional see L. W. Davis, "The Statistical Basis of Hydrodynamics and Kinetic Theory," (M. L. Report No. 755, Stanford University, 1960).

Next, using the representation Eq. (4-29) for a diagonal operator we have

$$n_d(x',-is) = n(x',-is) - \int_{-\infty}^{0} dt\, e^{\epsilon t}\, \dot{n}(x', t-is) \tag{4-39}$$

Inserting (4-39) into (4-38) and noting that

$$\mathrm{Tr}\, \rho_0 \vec{j} \equiv \langle \vec{j} \rangle_0 = 0 \tag{4-40}$$

$$\mathrm{Tr}\, \rho_0\, n(x',-is)\,\vec{j}(x) \equiv \langle n(x',-is)\,\vec{j}(x) \rangle_0 = 0 \tag{4-41}$$

we find

$$\langle \vec{j}(x) \rangle \doteq -\int_{-\infty}^{0} dt\, e^{\epsilon t} \int_0^{\beta} \frac{ds}{\beta} \int d^3x'\, \mu(x')\, \langle \dot{n}(x', t-is)\, \vec{j}(x) \rangle_0 \tag{4-42}$$

Eq. (4-40) states that the expectation value of the current flow in equilibrium vanishes. Eq. (4-41) follows as a consequence of time inversion symmetry. Under time inversion $\vec{j}$ changes sign but n does not. Using the continuity equation (4-32) we replace $\dot{n}$ by $-\vec{\nabla}\cdot\vec{j}$,

$$\langle \vec{j}(x) \rangle \doteq -\int_{-\infty}^{0} dt\, e^{\epsilon t} \int_0^{\beta} \frac{ds}{\beta} \int d^3x'\, \mu(x')\, \vec{\nabla}' \cdot \langle \vec{j}(x', t-is)\, \vec{j}(x) \rangle_0 \tag{4-43}$$

Integrating this by parts we obtain,

$$\langle \vec{j}(x) \rangle \doteq -\int_{-\infty}^{0} dt\, e^{\epsilon t} \int_0^{\beta} \frac{ds}{\beta} \int d^3x'\, \vec{\nabla}'\mu(x') \cdot \langle \vec{j}(x', t-is)\, \vec{j}(x) \rangle_0$$

$$+ \int_{-\infty}^{0} dt\, e^{\epsilon t} \int_0^{\beta} \frac{ds}{\beta} \oint d\vec{s}' \cdot \mu(x')\, \langle \vec{j}(x', t-is)\, \vec{j}(x) \rangle_0 \tag{4-44}$$

The correlation function $\langle \overline{J}(x',t-is)\overline{J}(x)\rangle_0$ decays to zero for $|\overline{x} - \overline{x}'|$ greater than some correlation length L_c. We expect L_c to be of order the mean free path of the particles in the lattice. When $\overline{x}$ is more than a correlation length away from the surface only the first term in Eq. (4-44) contributes. When $\overline{x}$ is within a correlation length of the surface, the surface term in Eq. (4-44) provides the correction to the current density due to the presence of the surface region.

For steady-state diffusion we expect

$$\nabla \cdot \langle \vec{j}(x) \rangle = 0 \tag{4-45}$$

when $\overline{x}$ is in a source free region (i.e. when $\overline{x}$ is not near the surface). Taking the gradient of the first term in Eq. (4-44) we have the volume integral

$$\int d^3x' \; \vec{\nabla}\mu(x') \cdot \langle \vec{j}(x', t-is)\, \vec{\nabla}\cdot\vec{j}(x) \rangle_0 \tag{4-46}$$

In component form this is

$$\int d^3x' \; \frac{\partial}{\partial x'_i}\mu(x') \frac{\partial}{\partial x_k} \langle j_i(x', t-is)\, j_k(x) \rangle_0 \tag{4-47}$$

For an isotropic medium the correlation function depends only on the distance $|\overline{x}' - \overline{x}|$, and orthogonal current densities are not correlated. In this case Eq. (4-47) becomes

$$\delta_{ik} \cdot \int d^3x' \; \frac{\partial}{\partial x'_i}\mu(x') \frac{\partial}{\partial x_i} \langle j_i(x', t-is)\, j_i(x) \rangle_0 \tag{4-48}$$

Integrating by parts and collecting components

$$\nabla\cdot\langle\vec{j}(x)\rangle \sim \int d^3x' \; \nabla'^2\mu(x') \, \langle j_i(x',t-is) \, j_i(x)\rangle_0 \tag{4-49}$$

The surface term vanishes since $\overline{x}$ is many correlation lengths from the surface. Therefore, the divergence of the current density will vanish inside the material if the chemical potential satisfies Laplace's equation,

$$\nabla^2\mu(x) = 0 \tag{4-50}$$

This is simply the statement that there are no sources or sinks of particles inside the material.

If the gradient of the chemical potential is constant over distance of order a correlation length then Eq. (4-44) becomes

$$\langle\vec{j}(x)\rangle \doteq -\vec{\nabla}\mu(x)\cdot\int_{-\infty}^{0} dt\, e^{\epsilon t}\int_0^{\beta}\frac{ds}{\beta}\int d^3x' \,\langle\vec{j}(x',t-is)\,\vec{j}(x)\rangle_0 \tag{4-51}$$

The local, diffusion transport coefficient is therefore[34]

$$D_{ik}(x) = \lim_{\epsilon\to 0}\int_{-\infty}^{0} dt\, e^{\epsilon t}\int_0^{\beta}\frac{ds}{\beta}\int d^3x' \,\langle j_k(x',t-is)\, j_i(x)\rangle_0 \tag{4-52}$$

This is similar to the transport coefficient relations obtained by R. Kubo et. al.[35]

---

34) As previously noted, for the case of a continuous spectrum, the limit $\epsilon\to 0$ is taken at the end of the calculation.

35) R. Kubo et. al., op. cit.

CHAPTER V

The Second Law of Thermodynamics (Reversible and Irreversible Processes)

The traditional statements of the Second Law of thermodynamics are associated with the names of Clausius and Kelvin.

> It is impossible to construct a cyclic process[36] whose only effect is to:
>
> 1) transfer heat from a heat reservoir at temperature T_1 to a heat reservoir at temperature T_2 where $T_2 > T_1$.
>
> 2) absorb heat from a heat reservoir and convert this heat into work.

Analysis of these statements implies the existence of a state function of the thermodynamic variables, the entropy S, and an absolute temperature scale T. These two thermodynamic quantities are related to the infinitesimal heat đQ absorbed by the system in a quasi-static process.

$$đQ = TdS \qquad (5\text{-}1)$$

A quasi-static process can be defined as an ordered succession of equilibrium states. A real process is a temporal succession of equilibrium and nonequilibrium states. For real processes the equality sign in (5-1) is replaced by an inequality.

$$đQ < TdS \qquad (5\text{-}2)$$

Definition of S_e and S_i

Experimentally the entropy difference between two equilibrium states is found by measuring and summing the increments

---

36) A cyclic process is one in which the system is eventually returned to its initial state.

$\frac{dQ}{T}$ which occur during a real process which connects the two states. This real process takes place over times long compared to the relaxation times characterizing the physical system. In this sense it approximates a quasi-static process, and we identify S_e with the symbol S appearing in the Second Law. Let the specific heat C for the particular process of interest be defined by

$$C = \frac{dQ}{dT} \tag{5-3}$$

Then the experimental entropy difference is

$$S_e(2) - S_e(1) = \int_1^2 \frac{dQ}{T} = \int_1^2 C \frac{dT}{T} \tag{5-4}$$

We have labeled the experimental entropy S_e to distinguish it from the information entropy S_i which will be introduced next.

The information theory entropy is defined as a functional of a density matrix ρ.

$$S_i[\rho] = - \mathrm{Tr}\, \rho \ln \rho \tag{5-5}$$

In a representation which diagonalizes ρ, Eq. (5-5) becomes

$$S_i = - \sum_n P_n \ln P_n \qquad P_n = \langle n | \rho | n \rangle$$

S_i is a consistent measure of the amount of uncertainty in the probability assignment of ρ. The uniqueness of this measure is discussed in Appendix A. In the case where ρ is separable, S_i is essentially the logarithm of the number of states in the HPM belonging to ρ.37) We have defined S_i as a dimension-

---

37) This corresponds to Boltzmann's original interpretation of log W as a measure of the number of experimentally compatible states.

less quantity.[38)]

The relationship between S_e and S_i can be inferred from the observation that the Gibbs' Canonical formalism predicts experimental results. A MEE consistent with the experimental expectation value of the energy gives the canonical density matrix of Gibbs. Therefore we conclude: in thermal equilibrium, the experimental entropy S_e, Eq. (5-4), forms an upper bound to the information theory entropy S_i, Eq. (5-5), for all possible density matrices compatible with the experimental expectation value of the energy.[39)] For thermal equilibrium in which S_e depends upon the values of additional variables $\langle F_i \rangle$, (e.g. mole numbers, volume, uniform rotational angular velocity) this same relationship between S_e and S_i holds. Let S_e depend upon the experimentally measured equilibrium values $\langle F_i \rangle$. Let $\bar{\rho}$ be the MEE density matrix consistent with these expectation values $\langle F_i \rangle$. Then

$$S_e = S_i[\bar{\rho}] \qquad (5\text{-}6)$$

This is an experimental fact since the MEE formalism reproduces the results of equilibrium statistical mechanics and these results agree with experiments. This relationship between S_e and S_i will now be used to define S_e for non-equilibrium states. This definition has three important consequences:

---

38) In discussing the relation between S_i and S_e a factor of k (Boltzmann's constant) will be needed. For convenience we shall temporarily work in the system of units where $k = 1$.

39) Implicit in this conclusion is the assumption that S_e and S_i are measured from the same reference state.

1) $S_e(t)$ is a function of experimentally measured quantities $\langle F_i(t) \rangle$; 2) for thermal equilibrium it reduces to the thermodynamic entropy; 3) $S_e(t)$ obeys the Second Law as we will now show.

The Time Development of the Entropy

The Second Law states that the experimental entropy S of an isolated system increases or remains the same. Now consider the time development of the information theory entropy, S_i. The Schrodinger equation implies that ρ develops in time under a unitary transformation. Since S_i is defined by the trace of a function depending only on ρ , it remains constant in time. This fact has seemed paradoxical when the distinction between S_e and S_i has been ignored.[40] Having distinguished between S_e and S_i the semantic origin of the paradox is obvious. Since S_e forms an upper bound to S_i, the fact that S_i is constant in time does not contradict the Second Law.

Before giving an analytic formulation of the Second Law, we shall sketch a geometrical argument due to Dr. E. T. Jaynes. This argument is based upon three fundamental notions: 1) the theorist must predict only experimentally reproducible results; 2) there exists a HPM of states associated with the MEE density matrix characterizing the experimentally allowed states; 3) the time evolution of the system is determined by the Schrodinger equation.[41]

---

40) Actually this is a sematic difficulty arising from the use of the word entropy to mean two different things.

41) Alternately, we could require the system to obey Hamilton's equations. The invariance of the dimensions of a manifold under a unitary transformation would be replaced by Liouville's theorem.

A MEE gives, for the initial density matrix, the broadest possible probability assignment consistent with the experimental information characterizing the initial state. The fact that experimental results are reproducible must therefore mean that almost all states consistent with the initial information give the same experimental results. These experimentally allowed states form the HPM associated with the initial density matrix. Let this manifold $\mathcal{M}(0)$ have dimension W(o). One of the properties of a separable density matrix is that

$$S_i[\rho_0] = -\mathrm{Tr}\,\rho_0 \ln \rho_0 = \ln W(0)$$

to order $N^{-\frac{1}{2}}$. N varies as the number of degrees of freedom of the system. Since the initial density matrix is obtained by a MEE consistent with the experimental constraints at time t_0, we have the equality

$$S_e(t_0) = S_i[\rho_0] = \ln W(0)$$

The Schrodinger equation implies that the system develops in time under a unitary transformation U (t, t_0).

$$i\dot{U} = \mathcal{H} U \qquad \hbar = 1$$

The initial manifold develops in time under U.

$$\mathcal{M}(t) = U(t, t_0)\,\mathcal{M}(0)$$

U represents a rotation of the state space which preserves the dimensions of the manifold. Therefore the manifold $\mathcal{M}(t)$ is of dimension W(o).

Since we are dealing with a system which gives experimentally reproducible results the manifold $\mathcal{M}(t)$ must be embedded in the manifold of states $\mathcal{M}_E(t)$ consistent with the experimental observations on the system at time t. Therefore the dimensionality $W(t)$ of the manifold $\mathcal{M}_E(t)$ must be greater than or equal to $W(0)$. However, $\mathcal{M}_E(t)$ forms the HPM of the MEE density matrix consistent with the experimental observations at time t. Therefore, from (5-6) we identify $S_e(t)$ with $\log W_E(t)$, where $W_E(t)$ is the dimensionality of $\mathcal{M}_E(t)$. This gives the Second Law.

$$S_e(t) = \log W_E(t) \geq \log W(t) = \log W(0) = S_e(t_0)$$

Analytic Formulation of the Second Law

Consider an experiment which measures the expectation values of certain dynamical variables $F_i (i = 1, \ldots N)$ belonging to a system.[42] At the initial time t_0 the expectation values are $\langle F_i \rangle_0$. The MEE density matrix consistent with this information is

$$\rho_0 = e^{-\lambda_0 - \sum_{i=1}^{N} \lambda_i F_i} \qquad (5\text{-}7)$$

$$\lambda_0 = \ln Z$$

$$\langle F_i \rangle_0 = -\frac{\partial}{\partial \lambda_i} \ln Z$$

$$Z = \mathrm{Tr}\, e^{-\sum_{i=1}^{N} \lambda_i F_i}$$

Following (5-6) we identify $S_e(t_0)$ with $S_1[\rho_0]$

$$S_e(t_0) = -\mathrm{Tr}\, \rho_0 \ln \rho_0 = \lambda_0 + \sum_{i=1}^{N} \lambda_i \langle F_i \rangle_0$$

---

[42] Throughout this discussion we shall assume that these measurements are of macroscopic quantities and do not produce important perturbations on the system.

The system develops in time under the Hamiltonian $\mathcal{H}(t)$.[43] The time dependence of $\mathcal{H}$ may be due to variations of external mechanical parameters such as electromagnetic field strengths. The equation of motion for an operator F_i is

$$\dot{F}_i(t) = i\,[\mathcal{H}(t), F_i(t)] \qquad F_i(t_0) = F_i$$

At a later time t the expectation values will be

$$\langle F_i(t)\rangle_0 = Tr\,\rho_0\,F_i(t)$$

The MEE density matrix consistent with these expectation values is

$$\rho_t = e^{-\lambda_0(t) - \sum_i \lambda_i(t) F_i} \tag{5-8}$$

with

$$\lambda_0(t) = \ln Z_t$$

$$\langle F_i(t)\rangle_0 = -\frac{\partial}{\partial\lambda_i(t)} \ln Z_t$$

$$Z_t = Tr\, e^{-\sum_i \lambda_i(t) F_i}$$

For reproducible experiments, the observed expectation values at time t, $\langle F_i \rangle_t$, will be equal to the calculated results $\langle F_i(t)\rangle_0$, if $\mathcal{H}(t)$ correctly describes the dynamics, and ρ_0 correctly describes the initial manifold of possible states. However, it is important to keep clearly in mind that the expectation values $\langle F_i(t)\rangle_0$, which determine ρ_t, are calculated numbers.

---

43) The system is considered as thermally isolated from its surroundings so that its time development is completely characterized by a Hamiltonian.

Identifying the experimental entropy $S_e(t)$ at time t with the information entropy of the MEE density matrix ρ_t we have

$$S_e(t) = -Tr\,\rho_t \ln \rho_t = \lambda_0(t) + \sum_i \lambda_i(t) \langle F_i(t) \rangle_0$$

This may be recast in the convenient form

$$S_e(t) = -Tr\,\rho(t) \ln \rho_t \tag{5-9}$$

where $\rho(t)$ is the time developed density matrix generated from ρ_0.

$$i\dot{\rho}(t) = [\mathcal{H}(t), \rho(t)] \qquad \rho(0) = \rho_0 \tag{5-10}$$

Eq. (5-9) will be taken as the fundamental relation giving the time dependent entropy of a system.44) In thermal equilibrium it reduces to the standard Gibbs' form. It provides a natural extension of this for non-equilibrium situations.

The analytic formulation of the Second Law is obtained by considering the entropy difference,

$$\Delta S = S_e(t) - S_e(t_0) = -Tr\,\rho(t) \ln \rho_t + Tr\,\rho_0 \ln \rho_0$$

$$\Delta S = Tr\,\rho(t) (\ln \rho(t) - \ln \rho_t) \tag{5-11}$$

Choose a representation $|n\rangle$ which diagonalizes $\rho(t)$. The concave property of the function $\ln x$ implies

$$\langle n | \ln \rho_t | n \rangle \leq \ln \langle n | \rho_t | n \rangle \tag{5-12}$$

---

44) H. Mori, Journ. Phys. Soc. (Japan), 11, 1029 (1956), has used this same definition for $S_e(t)$.

This inequality manifests itself as an increase in ΔS. It arises from the possibility that ρ_t and $\rho(t)$ do not commute.[45] Using (5-12) in (5-11) we find

$$\Delta S \geq \sum_m \langle m|\rho(t)|m\rangle \left(\ln \langle m|\rho(t)|m\rangle - \ln \langle m|\rho_t|m\rangle \right) \qquad (5\text{-}13)$$

$$\Delta S \geq \sum_m \langle m|\rho(t)|m\rangle \ln \left(\frac{\langle m|\rho(t)|m\rangle}{\langle m|\rho_t|m\rangle} \right)$$

Now using the inequality $\ln x \geq 1 - \frac{1}{x}$ it follows that

$$\Delta S \geq \sum_m \langle m|\rho(t)|m\rangle \left(1 - \frac{\langle m|\rho_t|m\rangle}{\langle m|\rho(t)|m\rangle} \right) \qquad (5\text{-}14)$$

The right hand side of this inequality vanishes since $\mathrm{Tr}\,\rho_t = \mathrm{Tr}\,\rho(t) = 1$. The final result is the Second Law.

$$\Delta S = S_e(t) - S_e(t_0) \geq 0$$

Reversible and Irreversible Processes

The dynamical equations which determine the time development of a system are invariant under time reversal. Classically this means that if all velocities[46] are reversed, and time proceeds in the reverse direction, the system will move back along its past trajectory in phase space. This is a classically allowed symmetry property since F = ma involves only second derivatives with respect to time.

---

[45] The entropy increase corresponding to the inequality (5-12) arises from the statistical features inherent in quantum mechanics. R. C. Tolman, op. cit., refers to (5-13) as Klein's Lemma. The important point is that for macroscopic systems the inequality in (5-13) is negligible. The significant inequality Eq. (5-14) arises from $\ln x \geq 1 - x^{-1}$.

[46] All external magnetic fields must also be reversed in direction.

In quantum mechanics the state of the system is represented by a wave function $\psi(t)$. Let $F(\overline{x},\overline{p})$ be a Hermitian operator depending upon position and momentum. The time reversed state function $\varphi(t)$ related to the state $\psi(t)$ may be characterized as follows.

$$F_\varphi(\vec{x},\vec{p},t) = \langle\varphi(t)|F(\vec{x},\vec{p})|\varphi(t)\rangle = \langle\psi(-t)|F(\vec{x},\vec{p})|\psi(-t)\rangle = F_\psi(\vec{x},-\vec{p},-t) \qquad (5\text{-}15)$$

We shall now give an explicit form for $\varphi(t)$ and show that if $\psi(t)$ is a possible solution of the Schrodinger equation then so is $\varphi(t)$. This is the sense in which the dynamics are invariant under time reversal. In this analysis we shall assume ψ is a scalar47) and H is an even function of the momentum.48)

Introduce a time inversion operator T.

$$\varphi(t) = T\,\psi(-t) \qquad (5\text{-}16)$$

Eq. (5-15) will be satisfied if T corresponds to the operation of taking the complex conjugate. T is therefore an antiunitary operator.49) Some useful properties of T are

$$T^{-1} = T$$

---

47) If ψ is a spinor then the time reversal operator mixes its components. It essentially changes spin-up to spin-down. See E. P. Wigner, Group Theory (Academic Press, New York, 1959), Chapter 26.

48) If a magnetic field is present the Hamiltonian will have an odd term in the momentum. However the basic time invariance symmetry Eq. (5-17) is maintained since the external magnetic fields must be reversed in direction under time reversal.

49) Let some state φ be expanded in terms of a complete, orthogonal set of functions u_n, $\varphi = \Sigma\, a_n u_n$. O is an antiunitary operator if for any set a_1, a_2, ... $O\varphi = \Sigma\, a_n^* O u_n$.

$$\langle T\psi_1 | \psi_2 \rangle = \langle \psi_1 | T\psi_2 \rangle^*$$

$$T^{-1} \vec{p}\, T = -\vec{p}$$

The last relation follows simply by considering $\overline{p}$ in the x-representation. Using these relations we shall now show that Eq. (5-15) is satisfied by (5-16).

$$\begin{aligned} F_\varphi(\vec{x}, \vec{p}, t) &= \langle \varphi(t) | F(\vec{x}, \vec{p}) | \varphi(t) \rangle \\ &= \langle T\psi(-t) | F(\vec{x}, \vec{p}) | T\psi(-t) \rangle \\ &= \langle \psi(-t) | T^{-1} F(\vec{x}, \vec{p})\, T | \psi(-t) \rangle^* \\ &= \langle \psi(-t) | F(\vec{x}, -\vec{p}) | \psi(-t) \rangle^* \\ &= \langle \psi(-t) | F(\vec{x}, -\vec{p}) | \psi(-t) \rangle = F_\psi(\vec{x}, -\vec{p}, -t) \end{aligned}$$

The last step follows since F is Hermitian.

If $\psi(t)$ satisfies the Schrodinger equation then so will the time reversed solution $\varphi(t)$.

$$i\hbar \frac{\partial}{\partial t} \psi(t) = H \psi(t)$$

To show this, let T operate on both sides of this equation. Since H is an even function of P it is invariant under the time reversal operator T.

$$T H T^{-1} = H$$

Therefore,

$$i\hbar \frac{\partial}{\partial t} T \psi(t) = H T \psi(t)$$

Changing t to -t one obtains the desired result,

$$i\hbar \frac{\partial}{\partial t} \phi(t) = H\phi(t)$$

This dynamic reversibility of the equations of motion is in contrast to the experimentally observed tendency for macroscopic processes to proceed spontaneously in only one direction. This latter behavior of macroscopic systems is referred to as irreversible. Just as in the case of the word entropy the word irreversible has lead to semantic difficulties. These difficulties have at times obscured the physics. Thus we find Tolman50) stating:

> "We thus find that the principle of dynamical reversibility would hold in the quantum mechanics in much the same way as in the classical mechanics. Hence, the introduction of the quantum mechanics--at least in its present form--cannot be regarded as throwing any new kind of light on the problem of the actual phenomenological irreversibility of thermodynamic processes. Just as in the classical mechanics, this irreversibility will have to be explained by considering the probable behavior of a collection or ensemble of systems rather than from consideration of the purely mechanical behavior of a single system."

Experiments are not performed on an ensemble of systems, but upon a single system. The fact that a process is observed to go in a certain direction must imply that this is true for the great majority of global quantum states consistent with the experimental control.

A fundamental criterion for irreversible processes is that formulated by Einstein*. According to Einstein, irreversibility arises from a lack of experimental control. An

---

50) R. C. Tolman, op. cit., Chapter XI, the last paragraph of Section 95.

* W. Ritz and A. Einstein, Physik. Zeits. 10, 323 (1909).

irreversible process can be made to proceed reproducibly in one direction in the laboratory, with a degree of control over initial conditions which fixes the values of only a few macroscopic parameters. The inverse process is not prohibited by any law of physics. However, in order to realize it reproducibly we would require a far greater degree of control, involving microscopic details of the initial state. In other words, the initial state for the inverse process has a very low a-priori probability, as long as we are restricted to macroscopic accuracy in controlling experimental conditions.

This same criterion for the existence of experimentally observed irreversible processes is found by examining the relation for $S_e(t)$, Eq. (5-9). The choice of the dynamical variables F_i whose expectation values at time t are calculated to specify ρ_t depends upon the specific experiment. As we increase this number we further restrict the form of the MEE density matrix ρ_t. If it were possible to make $\rho_t = \rho(t)$ then the entropy Eq. (5-9) would remain constant. However, this exact equality would require that the operators F_i form a complete, commuting set. Experimentally we do not have this microscopic control available.[51]

From this vantage point it is clear that the entropy Eq. (5-9) will increase in time. It is not clear that there exist isentropic experimental processes in which this entropy increase can approach zero. This seems to require the

---

[51] A complete measurement of this type would violate our assumption that the act of measurement does not perturb the system in an observable manner.

equality $\rho_t = \rho(t)$ which is experimentally unattainable. How then are we to explain the existence of isentropic processes?

The concept of a HPM provides an answer to this problem. In Chapter III we showed that under certain physically realized conditions a sharply defined manifold of states could be associated with a MEE density matrix. The sum of the probabilities associated with the states in this manifold approaches arbitrarily close to one. The probability associated with almost all states in this HPM is equal to W^{-1} where W represents the dimensions of the HPM. Furthermore, the information theory entropy associated with the MEE density matrix is equal to the logarithm of W.

In terms of these concepts we may now give a simple characterization of an experimentally reversible process. A reversible process is one in which sufficient control can be maintained so that the dimensionality of the HPM associated with ρ_t is equal to that of the HPM associated with ρ_o. The HPM of ρ_t must, of course, span the same manifold of states that the HPM of $\rho(t)$ spans. However, this last requirement is a consequence of the fact that we analyze only reproducible experiments. The possibility of isentropic experimental processes is thus directly associated with the existence of a HPM.

A more detailed understanding of the relationship between isentropic processes and HPM's is obtained by examining the matrix elements of the density matrix. The existence of HPM's implies that the only important matrix elements of ρ_t

and $\rho(t)$ are given by

$$\langle n|\rho(t)|n\rangle = \frac{1}{W} \qquad \text{for } |n\rangle \text{ in the HPM of } \rho(t)$$

$$\langle n'|\rho_t|n'\rangle = \frac{1}{W'} \qquad \text{for } |n'\rangle \text{ in the HPM of } \rho_t$$

If W' equals W, and the HPM's overlap then

$$\langle n|\rho(t)|n\rangle \left(\ln\langle n|\rho(t)|n\rangle - \ln\langle n|\rho_t|n\rangle\right) = 0$$

over the only region in which the matrix elements, appearing in the expression (5-13), give important contributions. Outside this HPM we do not make any requirements. Furthermore, we do not make any phase requirements on off diagonal matrix elements. As previously noted, the requirement that the HPM's overlap is always satisfied for reproducible experiments. Therefore the condition that a process be isentropic is simply W = W'.52)

The requirement $\rho_t = \rho(t)$ would indeed have meant that microscopic control would be necessary to obtain an isentropic process according to our definition of $S_e(t)$. However, the much weaker requirement upon the dimensionality of ρ_t can be maintained under suitable experimental situations. One such isentropic process will be analyzed in the next section.

Example of an Isentropic Process

We shall again consider the spin system characterized by a Hamiltonian $\mathcal{H}_t$.

$$\mathcal{H}_t = -H(t)M + D$$

---

52) See the footnote number 45 concerning Klein's Lemma.

Let the system be initially in equilibrium

$$\rho_0 = \frac{e^{-\beta_0 \mathcal{H}_0}}{Z_0} \qquad \mathcal{H}_0 = -H_0 M + D$$

where β_0 is determined by

$$\langle \mathcal{H}_0 \rangle_0 \equiv \mathrm{Tr}\, \rho_0 \mathcal{H}_0 = -\frac{\partial}{\partial \beta_0} \ln Z(\beta_0) \tag{5-17}$$

Now let the magnetic field change in time. According to Eq. (5-9) the entropy at time t is

$$S(t) = -\mathrm{Tr}\, \rho(t) \ln \rho_t$$

where

$$\rho_t = \frac{e^{-\beta(t) \mathcal{H}_t}}{Z_t}$$

$$Z_t = \mathrm{Tr}\, e^{-\beta(t) \mathcal{H}_t} \tag{5-18}$$

$$\mathrm{Tr}\, \rho(t) \mathcal{H}_t = -\frac{\partial}{\partial \beta(t)} \ln Z_t \tag{5-19}$$

and ρ (t) is the time developed density matrix

$$\rho(t) = U(t)\, \rho_0\, U^{+}(t)$$

$$i \dot{U}(t) = \mathcal{H}_t U(t) \qquad U(0) = 1 \qquad \hbar = 1$$

It is convenient to shift the time dependence to the dynamic variables and write Eq. (5-19) as

$$\langle \mathcal{H}(t) \rangle_0 \equiv \mathrm{Tr}\, \rho_0 \left(-H(t) M(t) + D(t)\right) = -\frac{\partial}{\partial \beta(t)} \ln Z_t \tag{5-20}$$

where $M(t) = U^{+}(t)\, M\, U(t) \qquad D(t) = U^{+}(t)\, D\, U(t)$

Using this notation

$$S(t) = \ln Z_t + \beta(t) \langle \mathcal{H}(t) \rangle_0$$

Taking the time derivative we find

$$\frac{d}{dt} S(t) = \frac{d}{dt} \ln Z_t + \dot{\beta}(t) \langle \mathcal{H}(t) \rangle_0 + \beta(t) \frac{d}{dt} \langle \mathcal{H}(t) \rangle_0 \qquad (5\text{-}21)$$

From Eq. (5-18)

$$\frac{d}{dt} \ln Z_t = -\dot{\beta}(t) \langle \mathcal{H}(t) \rangle_0 + \beta(t)\, \dot{H}(t) \langle M \rangle_t \qquad (5\text{-}22)$$

where

$$\langle M \rangle_t \equiv \mathrm{Tr}\, \rho_t\, M$$

Furthermore the change in the internal energy of the spin system is related to the work done by the external field.

$$d \langle \mathcal{H}(t) \rangle_0 = - \langle M(t) \rangle_0\, dH$$

Therefore

$$\frac{d}{dt} \langle \mathcal{H}(t) \rangle_0 = - \dot{H}(t) \langle M(t) \rangle_0 \qquad (5\text{-}23)$$

Using (5-22) and (5-23) in Eq. (5-21) we have

$$\frac{d}{dt} S(t) = \beta(t)\, \dot{H}(t) \left(\langle M \rangle_t - \langle M(t) \rangle_0 \right) \qquad (5\text{-}24)$$

First consider the case in which $\dot{H} < 0$ over a field range H_1 to H_2 ($H_1 > H_2$). Furthermore assume $H_2 \gg H_L$ where H_L is the local dipolar field defined by

$$H_L^2 = \frac{\mathrm{Tr}\, D^2}{\mathrm{Tr}\, M^2}$$

Let $\tau(H)$ be the relaxation time characterizing the exchange of Zeeman and dipolar energy.[53] Then we may obtain the following estimate for the right hand side of Eq. (5-24),

$$\langle M\rangle_t - \langle M(t)\rangle_0 \approx \frac{1}{H(t)} \tau \dot{H}(t) \langle M\rangle_t \tag{5-25}$$

The time rate of change of the entropy is

$$\frac{d}{dt} S(t) \approx \frac{\beta(t)\, \tau\, \dot{H}^2(t)\, \langle M\rangle_t}{H(t)}$$

Furthermore, in the high temperature approximation

$$S(t) \approx N \log(2I+1) - \frac{\beta(t) H(t)}{2} \langle M\rangle_t \tag{5-26}$$

Therefore

$$\frac{dS}{S_M - S} \approx \frac{2\tau\dot{H}}{H} \frac{dH}{H} \tag{5-27}$$

where

$$S_M \equiv N \log(2I+1)$$

From (5-27) the condition that the process be isentropic is

$$\frac{\tau\dot{H}}{H} << 1 \tag{5-28}$$

For the case in which $H \lesssim H_L$, Eq. (5-25) and Eq. (5-26) must be modified.

---

[53] N. Bloembergen, S. Shapiro, P. S. Pershan, J. O. Artman, Phys. Rev. 114, 445 (1959) have calculated $\tau(H)$. Within their approximation

$$\tau(H) \sim \exp\left[\frac{g^2\beta^2H^2}{2k^2\Delta}\right]$$

where for $H >> H_L$, Δ is the line width due to the non-secular terms in D. For $H \lesssim H_L$ they find $\tau(H)$ approaches the spin-spin relaxation time and is independent of H.

$$\langle M\rangle_t - \langle M(t)\rangle_0 \approx \frac{\tau \dot{H}(t) \langle M\rangle_t}{[H^2(t) + H_L^2]^{1/2}}$$

$$S(t) \approx N \log(2I+1) - \frac{\beta(t)}{2} \frac{[H^2(t) + H_L^2]}{H(t)} \langle M\rangle_t$$

The expression for the entropy change becomes

$$\frac{dS}{S_M - S} \approx \frac{2\tau \dot{H} H}{(H^2 + H_L^2)^{3/2}} dH \tag{5-29}$$

For H $\lesssim$ H_L, τ approaches the spin-spin relaxation time and is independent of H. Assuming H is constant we integrate Eq. (5-29).

$$\ln\left(\frac{S_M - S(2)}{S_M - S(1)}\right) \approx 2\tau\dot{H}\left[\frac{1}{(H^2(2) + H_L^2)^{1/2}} - \frac{1}{(H^2(1) + H_L^2)^{1/2}}\right]$$

In the low field range the isentropic condition is therefore

$$\frac{\dot{H}\tau}{H_L} << 1 \tag{5-30}$$

CHAPTER VI

The Dynamics of Irreversible Processes

Introduction

In this chapter we investigate certain aspects of the dynamics of irreversible processes. One aspect of irreversibility is the tendency of macroscopic systems to approach a state of equilibrium. We define equilibrium operationally by the requirement that the values of macroscopically measured quantities shall not change in time. More specific types of equilibrium are further characterized by the requirement that different parts of the system have the same value of some intensive property (e.g. temperature, thermal equilibrium; pressure, mechanical equilibrium; chemical potential, chemical equilibrium). One of the fundamental problems of irreversibility is the determination of the properties of the Hamiltonian $\mathcal{H}$, the initial state, and the observed dynamic variables F_i which are necessary and sufficient to imply that the system approaches equilibrium.

A second and related aspect of irreversibility is the time decay of thermal equilibrium fluctuation correlations. Let F be a dynamic variable whose time development is determined by the Heisenberg equation of motion.

$$\dot{F}(t) = i\,[\mathcal{H}, F(t)] \qquad (\hbar = 1)$$

Let ρ represent the equilibrium density matrix for the system. An autocorrelation function can be defined by the relation

$$C(t) = \frac{Tr\,\rho F F(t)}{Tr\,\rho F^2} = \frac{\langle F F(t)\rangle}{\langle F^2\rangle} \tag{6-1}$$

where F stands for F(0). For a macroscopic system we expect C(t) to vanish for large t.[54] Actually C(t) should become very small for t long compared to some characteristic time τ_c called the correlation time of F.

These two aspects of irreversibility are closely related. The time dependence of a dynamic variable can be related to correlation functions of the interaction responsible for its time development. In the next section we shall discuss this relationship. We will then examine some of the properties of the correlation functions when the energy spectrum of the system is discrete. As an example of this case we evaluate an autocorrelation function for a ring of N exchange-coupled spins. We then examine the dependence of this correlation function on N and t. For finite N the time dependence is almost-periodic.[55] However if we take the limit $N \to \infty$ before

---

54) In general C(t) should approach a constant value for large t. If $[\mathcal{H},\rho] = 0$ and $Tr\rho F = 0$, then we expect the correlation function to vanish eventually. If $Tr\rho F = \langle F\rangle \neq 0$, then consider the correlation function of the operator $F - \langle F\rangle$. It is primarily this class of thermal equilibrium fluctuation correlation functions which we shall be interested in.

55) A set of real numbers $\{\tau_i\}$ is called relatively dense if there exists a number T such that every interval $t < \tau < t + T$ of length T contains at least one member of the set. A continuous function f(t) is called almost-periodic if for any $\epsilon > 0$ there exists a relatively dense set of numbers $\{\tau_i\}$ such that $|f(t + \tau_i) - f(t)| \leq \epsilon$ for $-\infty < t < \infty$. See Harald Bohr, Almost-Periodic Functions, (Chelsea Publishing Co., New York (1947)).

the limit $t \to \infty$, the autocorrelation function goes to zero in the latter limit. In the limit $N \to \infty$ the energy spectrum of the system becomes continuous. This is responsible for the qualitative difference between the asymptotic time behavior of a finite and an infinite system.[56)]

The properties of the autocorrelation function are then examined for the case of a continuous energy spectrum. As an example of this case the autocorrelation function $\langle E_x E_x(t,\bar{x})\rangle$ is calculated for a thermal equilibrium radiation field. E_x stands for $E_x(0,0)$. We conclude by applying some of the results obtained from our analysis of the time behavior of correlation functions to the fundamental problem of the approach to equilibrium of macroscopic observables.

Some Examples of the Relationship Between Correlation Functions and the Time Development of Dynamic Variables

The Green's functions, or propagators, associated with a many body system have the form of correlation functions. This is the fundamental relationship between the time evolution of a system and the correlation function. Rather than review this formalism[57)] we will present some examples which illustrate this relationship.

---

56) This is associated with the Poincaré recurrence phenomenon of a finite system. Poincaré recurrence means that a given initial state will recur, not exactly, but to any desired degree of accuracy, infinitely often. Classically the initial state is specified by a point in phase space. Quantum mechanically it can be specified by the expectation values of a complete set of commuting observables.

57) P. C. Martin and J. Schwinger, Phys. Rev. 115, 1342 (1959)
L. D. Landau, JETP 34, 262 (1958).

Consider a system of interacting spins having a Hamiltonian

$$\mathcal{H}_0 = -H_0 M_z + D$$

Let D represent, for instance, the dipole-dipole interaction. Suppose the initial state of the system is characterized by an average value of energy, and an external magnetic field $H_o + \Delta H$. At $t = 0$ let ΔH vanish. What is the time development of the Z-component of magnetization?

$$\langle M_z(t)\rangle = Tr\,\rho\, M_z(t) \tag{6-2}$$

where

$$\rho = \frac{e^{-\beta(\mathcal{H}_0 - \Delta H M_z)}}{Z(\beta)}$$

$$Z(\beta) = Tr\, e^{-\beta(\mathcal{H}_0 - \Delta H M_z)}$$

$$M_z(t) = e^{i\mathcal{H}_0 t} M_z\, e^{-i\mathcal{H}_0 t}$$

Expanding the density matrix we obtain in lowest order

$$\rho \doteq \frac{e^{-\beta\mathcal{H}_0}\left(1 - \Delta H \int_0^\beta ds\, M_z(-is)\right)}{Z_0(\beta)\left(1 - \beta\Delta H\langle M_z\rangle_0\right)} \tag{6-3}$$

where

$$M_z(-is) = e^{s\mathcal{H}_0} M_z\, e^{-s\mathcal{H}_0}$$

$$Z_0(\beta) = Tr\, e^{-\beta\mathcal{H}_0}$$

$$\langle M_z\rangle_0 = \frac{Tr\, e^{-\beta\mathcal{H}_0} M_z}{Z_0(\beta)}$$

Substituting (6-3) into (6-2) we obtain, in this order

$$\langle M_z(t)\rangle \doteq \langle M_z\rangle_0 - \Delta H \langle M_z\rangle_0^2 \int_0^\beta ds \left[\frac{\langle M_z(-is)\, M_z(t)\rangle_0}{\langle M_z\rangle_0^2} - 1 \right] \tag{6-4}$$

Since $\mathcal{H}_0$ commutes with $\exp(-\beta\mathcal{H}_0)$ we may write the expectation value $\langle M_z(-is)M_z(t)\rangle$ in the form of a correlation function[58]

$$\frac{\langle M_z\, M_z(t+is)\rangle_0}{\langle M_z\rangle_0^2}$$

The higher order corrections involve higher order correlation functions. For example the next order correction to Eq. (6-4) contains the third order correlation function

$$\langle M_z\, M_z(-is_2+is_1)\, M_z(t+is_1)\rangle_0$$

As a second example, consider the problem of the interaction of a magnetic moment with a thermal equilibrium radiation field.

$$\mathcal{H} = -\mu_0\, \bar{\sigma}\cdot\bar{H} + W$$

$$W = \sum_{\lambda k} \omega_k\, a^+_{k\lambda}\, a_{k\lambda} \qquad (\hbar = 1)$$

$$\bar{H} = i \sum_{\lambda k} \left(\frac{2\pi c^2}{\omega_k V}\right)^{1/2} (\bar{k}\times\bar{\epsilon}_{k\lambda})\left(a^+_{k\lambda}\, e^{-ik\cdot x} - a_{k\lambda}\, e^{ik\cdot x}\right)$$

$$[a_{k\lambda}, a^+_{k'\lambda'}] = \delta_{kk'}\,\delta_{\lambda\lambda'}$$

---

[58] Since Tr $\rho M_z \neq 0$ and $[\mathcal{H}_0, \rho] \neq 0$ we do not expect this correlation function to decay to zero. However, we do expect it to eventually approach an equilibrium value.

where we have quantized the radiation field in a box of volume V. Going to the field interaction representation we have

$$\dot{\sigma}_z = -i\mu_0 \bar{\vec{H}}(t) \cdot [\vec{\sigma}, \sigma_z] \tag{6-5}$$

where $\bar{H}(t) = e^{iwt}\bar{H}\, e^{-iwt}$. Computing the commutation relation in (6-5)

$$\dot{\sigma}_z(t) = 2\mu_0 (H_y(t)\sigma_x(t) - H_x(t)\sigma_y(t)) \tag{6-6}$$

The equations for $\sigma_x(t)$ and $\sigma_y(t)$ follow by cylic permutation

$$\dot{\sigma}_x(t) = 2\mu_0 (H_z(t)\sigma_y(t) - H_y(t)\sigma_z(t)) \tag{6-7}$$

$$\dot{\sigma}_y(t) = 2\mu_0 (H_x(t)\sigma_z(t) - H_z(t)\sigma_x(t)) \tag{6-8}$$

Integrating Eqs. (6-7) and (6-8) to obtain formal relations for $\sigma_x(t)$ and $\sigma_y(t)$ we write Eq. (6-6) as

$$\begin{aligned} \dot{\sigma}_z(t) &= 2\mu_0 (H_y(t)\sigma_x(0) - H_x(t)\sigma_y(0)) \\ &+ 4\mu_0^2 \int_0^t dt_1 \{ H_x(t) H_z(t_1)\sigma_x(t_1) - H_x(t)H_x(t_1)\sigma_z(t_1) \\ &\quad - H_y(t)H_y(t_1)\sigma_z(t_1) + H_y(t)H_z(t_1)\sigma_y(t_1) \} \end{aligned} \tag{6-9}$$

Assume that the initial state of the radiation field is described by the thermal equilibrium density matrix

$$\rho_W = \frac{e^{-\beta W}}{Z_W(\beta)} \qquad Z_W(\beta) = \mathrm{Tr}\, e^{-\beta W}$$

Multiplying Eq. (6-9) by ρ_W and tracing out the radiation field one obtains

$$\dot{\sigma}_z(t) \doteq -8\mu_0^2 \int_0^t dt_1 \, \langle H_x(t) H_x(t_1) \rangle_W \, \sigma_z(t_1) \qquad (6\text{-}10)$$

where $\langle 0 \rangle_W \equiv \mathrm{Tr}\rho_W 0$. The approximation made in obtaining (6-10) consists in the replacement

$$\langle H_x(t) H_x(t_1) \, \sigma_z(t_1) \rangle_W \longrightarrow \langle H_x(t) H_x(t_1) \rangle_W \, \sigma_z(t_1)$$

That is, we include only that part of the trace in which $H_x(t)$ emits or absorbs a quantum absorbed or emitted by $H_x(t_1)$. This is the lowest order correlation approximation.

In obtaining Eq. (6-10) we further note that

$$\langle H_x \rangle_W = \langle H_x(t) H_y(t_1) \rangle_W = 0$$

$$\langle H_x(t) H_x(t_1) \rangle_W = \langle H_y(t) H_y(t_1) \rangle_W$$

The time dependence of $\langle H_x(t) H_x(t_1) \rangle_W$ can be simplified since $\exp(iWt)$ commutes with ρ_W.

$$\langle H_x(t) H_x(t_1) \rangle_W = \langle H_x(t-t_1) \, H_x \rangle_W \equiv \langle H_x^2 \rangle_W \, f(t-t_1) \qquad (6\text{-}11)$$

Using this in Eq. (6-10) we have the lowest order correlation function relation

$$\dot{\sigma}_z(t) \doteq -8\mu_0^2 \langle H_x^2 \rangle_W \int_0^t dt_1 \, f(t-t_1) \, \sigma_z(t_1) \qquad (6\text{-}12)$$

The Laplace transform of Eq. (6-12) gives

$$s\,\bar{\sigma}_z(s) - \sigma_z(0) \doteq -8\mu_0^2 \langle H_x^2 \rangle_W \, \bar{f}(s) \, \bar{\sigma}_z(s)$$

where

$$\bar{\sigma}_z(s) = \int_0^\infty dt \, e^{-st} \sigma_z(t)$$

$$\bar{f}(s) = \int_0^\infty dt\, e^{-st} f(t)$$

Solving this equation one obtains

$$\bar{\sigma}_z(s) \doteq \frac{\sigma_z(0)}{s + 8\mu_0^2 \langle H_x^2 \rangle_W \bar{f}(s)} \tag{6-13}$$

Therefore in this approximation the time behavior of $\sigma_z(t)$ is determined by the zeros of

$$s - 8\mu_0^2 \langle H_x^2 \rangle_W \bar{f}(s) \tag{6-14}$$

For example if $f(t) = e^{-t/\tau}$

$$\bar{f}(s) = \frac{1}{s + \frac{1}{\tau}}$$

and Eq. (6-14) has zeros at

$$s = -\frac{1}{2\tau} \pm \sqrt{\left(\frac{1}{2\tau}\right)^2 - 8\mu_0^2 \langle H_x^2 \rangle_W}$$

Since there is no zero at s = 0, $\sigma_z(t)$ vanishes as t goes to infinity.

Later we shall explicitly evaluate f(t) for the black-body radiation field. It does not have the form of a simple exponential decay. Furthermore, using the correct $\bar{f}(s)$, Eq. (6-14) is found to have a zero at s = 0. Therefore, in this lowest order approximation $\sigma_z(t)$ remains finite as t goes to infinity.[59] Higher order approximations involving correlation functions of the form $\langle H_x(t)H_z(t_1)H_z(t_2)H_x(t_3)\rangle$ are responsible for the long time decay of $\sigma_z(t)$.

Correlation Functions for Systems Having Discrete Energy Spectra

In general we may define an n^th^ order correlation function,

$$\frac{\langle F_i(t_1)\, F_j(t_2) \cdots F_k(t_n)\rangle}{\langle F_i F_j \cdots F_k\rangle} \tag{6-15}$$

The expectation value may be taken for a single quantum state or as an average over a manifold of states. The time dependence

---

59) Consider the case in which the spin states are separated by some energy ϵ. Suppose the spin is initially in the upper state. The transition probability to the lower state is given in the lowest order Born approximation by

$$w = 2\pi \sum_{WW'} P(W)\, |\langle \downarrow W' | -\mu\cdot H | \uparrow W\rangle|^2\, \delta(E_W + \epsilon - E_{W'})$$

$$= 2\pi \langle n_k + 1\rangle \int \frac{d^3k}{(2\pi)^3}\, V\left(\frac{2\pi c^2}{\omega_k V}\right) \mu_0^2 k^2 (\cos^2\theta + 1)\, \delta(\epsilon - ck) \qquad \text{(a)}$$

The energy dependence is therefore

$$\int dk\, k^3\, \delta(\epsilon - ck) = \frac{\epsilon^3}{c^4}$$

and w vanishes as $\epsilon \to 0$. This is exactly the situation for a spin interacting with a radiation field (no external magnetic field). The exponential correlation function $\exp(-t/\tau)$, which implied relaxation, grossly misrepresents the low energy behavior of the radiation spectrum. This is the danger inherent in arbitrarily replacing a correlation function by an exponential decay.

The actual relaxation mechanism appears in higher Born approximations. Fluctuations in the black-body radiation field produce a fluctuating energy separation ϵ of order $\mu\langle H_z^2\rangle_W^{\frac{1}{2}}$. During these fluctuations, transitions can occur, which in lowest order produce a transition probability given by Eq. (a) in which $\epsilon = \mu\langle H_z^2\rangle_W^{\frac{1}{2}}$.

of the F_i are determined by the Heisenberg equations of motion

$$\dot{F}_i = i\,[\mathcal{H}, F_i] \qquad (\hbar = 1)$$

The autocorrelation function for a dynamic variable F is obtained as a special case of (6-15)

$$\frac{\langle F(t_1)\, F(t_2)\rangle}{\langle F^2\rangle} \tag{6-16}$$

If the expectation value is taken with respect to an energy eigenstate or as an average over a density matrix which commutes with the Hamiltonian, Eq. (6-16) is a function of $t_2 - t_1$.

$$C(t) = \frac{\langle F\, F(t)\rangle}{\langle F^2\rangle} \qquad t = t_2 - t_1 \tag{6-17}$$

Let the eigenstates of H have a discrete spectrum

$$H|n\rangle = \omega_n |n\rangle \tag{6-18}$$

where n stands for the complete set of quantum numbers needed to specify the state of the system. We shall assume that all of these quantum numbers are discrete. The normalization is then

$$\langle n | n'\rangle = \delta_{nn'}$$

where the δ-symbol represents a product of Kronecker δ's for all the quantum numbers needed to specify the state. Let C(t), Eq. (6-17) be defined with respect to a density matrix

$$C(t) = \frac{\mathrm{Tr}\,\rho\, F\, F(t)}{\mathrm{Tr}\,\rho\, F^2} \tag{6-18}$$

$$[\mathcal{H}, \rho] = 0$$

$$F(t) = e^{iHt} F e^{-iHt}$$

Taking the trace, Eq. (6-18), in the energy representation one obtains

$$C(t) = \frac{\sum_{nm} \langle n|\rho|n\rangle \, |\langle n|F|m\rangle|^2 e^{i\omega_{mn}t}}{\sum_{nm} \langle n|\rho|n\rangle \, |\langle n|F|m\rangle|^2} \tag{6-19}$$

where

$$\omega_{mn} = \omega_m - \omega_n$$

Take the Laplace transform of Eq. (6-19)

$$\int_0^\infty dt \, e^{-st} C(t) = \frac{\sum_{nm} \frac{\langle n|\rho|n\rangle \, |\langle n|F|m\rangle|^2}{s - i\omega_{mn}}}{\sum_{nm} \langle n|\rho|n\rangle \, |\langle n|F|m\rangle|^2}$$

This has poles along the imaginary axis of the s-plane.

$$s = i\omega_{mn}$$

Therefore, unless $\langle n|F|m\rangle = 0$ when $\omega_{mn} \neq 0$ the long time behavior of the correlation function is oscillatory.* More precisely, C(t) is an almost-periodic function. If $\langle n|F|m\rangle = 0$ when $\omega_{mn} \neq 0$, then the correlation function is identically equal to unity for all time.

The long time oscillatory behavior seems in direct contrast to the experimentally observed (irreversible) decay of the correlation function.[60] We may understand these two

---

[60] We consider thermal equilibrium fluctuation correlations.

* We assume $\langle n|\rho|n\rangle \neq 0$.

aspects of C(t) by considering Eq. (6-19). Initially at $t = 0$ all phases are zero and the matrix elements add up in a constructive manner to give $C(t) = 1$. This represents an upper bound on C(t) since

$$\sum_{mn} \langle m|\rho|m\rangle |\langle n|F|m\rangle|^2 e^{i\omega_{mn}t} \leq \sum_{mn} \langle m|\rho|m\rangle |\langle n|F|m\rangle|^2 \qquad (6\text{-}20)$$

As t increases, there occurs a dephasing of the various terms of the left hand sum, Eq. (6-20). It is this dephasing which causes the correlation function to decay initially. In fact, dephasing is the dynamic cause of irreversibility. For a sufficiently long time associated with the distribution of ω_{mn} values, C(t) will again approach arbitrarily close to one. Roughly, one may say that this time is given by the inverse frequency ω^{-1}; where ω is commensurable with the energy differences ω_{mn} of the dominant matrix elements occuring in the sum on the left hand side of Eq. (6-20). This Poincaré recurrence is a well known property of almost-periodic functions.61) We thus see the relationship between the possible time decay of correlations and the almost-periodic nature of their time dependence. According to Smoluchowski,62) an irreversible process is one whose initial macroscopic state has a recurrence time long compared to laboratory times. We now consider an illustration of these ideas.

---

61) Harald Bohr, op. cit.

62) M. von Smoluchowski, Physik. Z. 13, 1069 (1912); 14, 261 (1913).

Example of a Correlation Function for a Ring of N Exchange-Coupled Spins

Consider a ring of N exchange-coupled spins. The Hamiltonian is

$$\mathcal{H} = \hbar\omega \sum_{\mu=1}^{N} I_{\mu,\mu+1} \tag{6-21}$$

$$I_{\mu,\mu+1} = \frac{1 + \vec{\sigma}_{\mu} \cdot \vec{\sigma}_{\mu+1}}{2}$$

Initially all spins are spin-down except for the k^th^ one which is spin-up. We shall study the time evolution of this state. In particular we evaluate the spin-up correlation function for the k^th^ spin

$$C(t) = \langle \psi_k | \sigma_k^+ \sigma_k^- \sigma_k^+(t) \sigma_k^-(t) | \psi_k \rangle \tag{6-22}$$

The expectation value is taken with respect to the initial state

$$\psi_k = \beta_1 \beta_2 \cdots \beta_{k-1} \alpha_k \beta_{k+1} \cdots \beta_N \tag{6-23}$$

where β represents a spin-down state and α represents a spin-up state. The operator $(\sigma_k^+)(\sigma_k^-)$ gives 1 if the k^th^ spin is up and 0 if it is down. Let the operator $(\sigma_k^+)(\sigma_k^-)$ operate to the left in (6-22) to give

$$C(t) = \langle \psi_k | \sigma_k^+(t) \sigma_k^-(t) | \psi_k \rangle$$

$$= \langle \psi_k | U^+(t) \sigma_k^+ \sigma_k^- U(t) | \psi_k \rangle$$

and

$$C(t) = \langle \psi_k(t) | \sigma_k^+ \sigma_k^- | \psi_k(t) \rangle \tag{6-24}$$

where $U(t) = \exp(-\frac{i\mathcal{H}t}{\hbar})$ and

$$\psi_k(t) = U(t)\,\psi_k$$

We shall now calculate $\psi_k(t)$

Let

$$\Sigma_z = \sum_{n=1}^{N} \sigma_{z\,n} \qquad (6\text{-}25)$$

This is a constant of the motion

$$[\Sigma_z, \mathcal{H}] = 0$$

Therefore, the time development of an eigenstate of Σ_z takes place only in the manifold of states with the same Σ_z eigenvalue. The manifold with eigenvalue $\Sigma_z' = 1 - N$ is spanned by the basis

$$\begin{aligned} \psi_1 &= \alpha_1 \beta_2 \cdots \beta_N \qquad (6\text{-}26) \\ \psi_2 &= \beta_1 \alpha_2 \beta_3 \cdots \beta_N \\ &\cdots \\ \psi_N &= \beta_1 \beta_2 \cdots \beta_{N-1} \alpha_N \end{aligned}$$

From this basis we construct another orthogonal basis,

$$\Phi_\nu = \frac{1}{N^{1/2}} \sum_{n=1}^{N} e^{-i\frac{2\pi}{N}\nu n} P_n \psi_1 \qquad \nu = 1, 2, \cdots, N \qquad (6\text{-}27)$$

$$P_n \psi_1 \equiv \psi_n$$

For this basis we have,

$$\mathcal{H}\Phi_\nu = \left[\hbar\omega(N-2) + 2\hbar\omega\cos\tfrac{2\pi}{N}\nu\right]\Phi_\nu \qquad (6\text{-}28)$$

Therefore, Ψ_ν are eigenstates of energy with eigenvalues[63]

$$E_\nu = \hbar\omega(N-2) + 2\hbar\omega \cos\frac{2\pi\nu}{N} \tag{6-29}$$

$$\nu = 1, 2, \cdots, N$$

The initial state of the system is assumed to be,

$$\phi_K = \beta_1 \beta_2 \cdots \beta_{K-1} \alpha_K \beta_{K+1} \cdots \beta_N$$

In order to find the time development of this state we expand it in terms of the Ψ_ν basis.

$$\phi_K = \frac{1}{N^{1/2}} \sum_{\nu=1}^{N} e^{i\frac{2\pi\nu K}{N}} \Psi_\nu$$

Its time development is given by

$$\phi_K(t) = e^{-\frac{iHt}{\hbar}} \phi_K = \frac{1}{N^{1/2}} \sum_{\nu=1}^{N} e^{i\left(\frac{2\pi\nu K}{N} - \frac{E_\nu t}{\hbar}\right)} \Psi_\nu \tag{6-30}$$

Now we evaluate the expectation value of $(\sigma_K^+)(\sigma_K^-)$, Eq. (6-24).

$$C(t) = \frac{1}{N^2} \sum_{\nu\mu m l=1}^{N} e^{i\frac{2\pi}{N}[\nu(K-m)+\mu(l-K)] + \frac{iE_\nu t}{\hbar} - i\frac{E_\mu t}{\hbar}} \langle P_m \phi_1 | \sigma_K^+ \sigma_K^- | P_l \phi_1 \rangle$$

$$= \frac{1}{N^2} \sum_{\nu,\mu=1}^{N} e^{i\frac{(E_\nu - E_\mu)t}{\hbar}} \tag{6-31}$$

since

$$\langle P_m \phi_1 | \sigma_K^+ \sigma_K^- | P_l \phi_1 \rangle = \delta_{ml}\, \delta_{mK}$$

For t = 0 we have as expected

$$C(0) = 1$$

---

63) This solution was first given by H. Bethe, Zs. f. Phys. <u>71</u>, 205 (1931).

From Eq. (6-29) it follows that

$$E_\nu - E_\mu = 2\hbar\omega\left(\cos\frac{2\pi}{N}\nu - \cos\frac{2\pi}{N}\mu\right)$$

Inserting this into Eq. (6-31) we proceed to calculate

$$C(t) = \frac{1}{N^2}\sum_{\nu,\mu=1}^{N} e^{i2\omega t\left(\cos\frac{2\pi}{N}\nu - \cos\frac{2\pi}{N}\mu\right)}$$

$$= \frac{1}{N^2}\sum_{\nu,\mu=1}^{N}\ \sum_{n,m=-\infty}^{\infty} e^{in\left(\frac{2\pi\nu}{N}+\frac{\pi}{2}\right)} J_n(2\omega t)\, e^{-im\left(\frac{2\pi}{N}\mu+\frac{\pi}{2}\right)} J_m(2\omega t)$$

Using the relation

$$\frac{1}{N}\sum_{\nu=1}^{N} e^{i\frac{m2\pi\nu}{N}} = \sum_{\lambda=-\infty}^{\infty} \delta_{m,\lambda N}$$

we obtain[64]

$$C(t) = |f(t)|^2 \tag{6-32}$$

$$f(t) = \sum_{\lambda=-\infty}^{\infty} e^{i\frac{N\pi\lambda}{2}} J_{\lambda N}(2\omega t) \tag{6-33}$$

For a finite number of spins, C(t) is an almost-periodic function. Specifically, for the case N = 2 we find

$$f(t) = \sum_{\lambda=-\infty}^{\infty} e^{i\pi\lambda} J_{2\lambda}(2\omega t) = \cos 2\omega t$$

and

$$C(t) = \cos^2(2\omega t)$$

Now consider the limiting case for N ⟶ ∞

$$\lim_{N\to\infty} f(t) = J_0(2\omega t)$$

---

[64] It is interesting to note that E. Schrodinger found a similar Bessel function relation for a ring composed of N equal mass points coupled by identical springs. Ann. d. Phys. 44, 1916 (1914).

and the correlation function has an irreversible, oscillatory decay to zero.

$$C(t) = J_0^2(2\omega t) \tag{6-34}$$

In this limit the energy spectrum has become continuous. In the next section we shall examine correlation functions for systems having continuous energy spectra. The possibility of an asymptotic time decay is directly associated with the existence of a continuous spectrum.

For large, but finite N we shall examine the behavior of f as a function of Z = 2ωt.

$$f(z) = J_0(z) + e^{\frac{iN\pi}{2}} J_N(z) + e^{-\frac{iN\pi}{2}} J_{-N}(z)$$

$$+ e^{iN\pi} J_{2N}(z) + e^{-iN\pi} J_{-2N}(z) + \cdots$$

Assume N is even,

$$f(z) = J_0(z) + 2\left(J_N(z) + J_{2N}(z) + \cdots \right)$$

For large values of N and Z we have the asymptotic relations,[65]

$$z < N \qquad J_N(z) \sim \frac{e^{N(\tanh\alpha - \alpha)}}{\sqrt{2\pi N \tanh\alpha}} \qquad \tanh\alpha = \sqrt{1 - \left(\frac{z}{N}\right)^2}$$

$$z \sim N \qquad J_N(z) \sim \frac{\Gamma\left(\frac{1}{3}\right)}{2^{2/3}\, 3^{\frac{1}{6}}\, \pi N^{1/3}}$$

---

65) P. M. Morse and H. Feshbach, <u>Methods of Theoretical Physics</u>, Vol. I, (McGraw-Hill Book Company, Inc., New York, 1953) p. 631.

$$z > N \quad J_N(z) \sim \sqrt{\frac{2}{\pi N \tan\beta}} \cos\left[N\tan\beta - N\beta - \frac{\pi}{4}\right] \qquad \tan\beta = \sqrt{\left(\frac{z}{N}\right)^2 - 1}$$

Therefore, for $Z < N$

$$f(z) - J_0(z) \sim \frac{e^{-\gamma N}}{N^{1/2}} \qquad \gamma > 0$$

For $Z \sim N$

$$f(z) - J_0(z) \sim \frac{1}{N^{1/3}}$$

In order for the infinite series $J_N(Z)+J_{2N}(Z)+\ldots$ to give a contribution of order 1 we must therefore assume $Z > N$. In this case we have

$$J_N(z) \sim \frac{1}{z^{1/2}} \cos\left[Nz - \frac{N\pi}{2} - \frac{\pi}{4}\right]$$

We will need at least $Z^{\frac{1}{2}}$ terms like this to obtain a contribution of order one. We shall require that Z be greater than the order of the $Z^{\frac{1}{2}\text{th}}$ term. This implies

$$z^{1/2} N < z \implies z > N^2$$

Therefore the time τ_p which must elapse before f(t) departs appreciably from $J_0(2\omega t)$ is given by,

$$\tau_p > \frac{N^2}{\omega} \tag{6-35}$$

Finally, we consider an alternative method of estimating this Poincaré recurrence time. It gives essentially the same result as Eq. (6-35). Furthermore it shows the relationship between τ_p and the inverse frequency commensurable with a majority of the energy differences in the sum, Eq. (6-31).

For large N the energy level separation becomes

$$E_\nu - E_{\nu+1} = 2\hbar\omega\left(\cos\tfrac{2\pi}{N}\nu - \cos\tfrac{2\pi}{N}(\nu+1)\right) \doteq 2\hbar\omega\left(\tfrac{2\pi}{N}\right)^2\nu$$

A large number[66] of the energy differences appearing in the sum

$$\frac{1}{N^2}\sum_{\nu,\mu=1}^{N} e^{\frac{i(E_\nu - E_\mu)t}{\hbar}}$$

are very nearly multiples of

$$2\hbar\omega\left(\frac{2\pi}{N}\right)^2$$

Therefore the Poincaré recurrence time for C(t) is given by

$$\tau_p \sim \frac{N^2}{4\pi\omega} \tag{6-36}$$

For $\hbar\omega \sim 1$e.v. and $N \sim 10^{23}$ we have

$$\tau_p \sim \frac{10^{46}}{10^{22}} = 10^{22}\text{sec} \sim 10^{15}\text{ years}$$

Actually, this is short as far as most Poincaré recurrence times for systems with 10^{23} particles.[67] This is a consequence of the relatively simple structure of $\mathcal{H}$ (i.e. it only couples nearest neighbors) and the special choice of the initial state having $\Sigma' = N - 1$. Since Σ' is a constant of the motion, when the k^th spin is spin-up the remaining 10^{23}-1 spins are automatically spin-down.

---

[66] Large means of order αN for $\alpha \neq 0$.

[67] For multiply periodic Hamilton-Jacobi systems, H. Frisch shows that $\tau_p \sim \epsilon^{-N}$. ϵ is the error of recurrence and N is the number of degrees of freedom of the system. Phys. Rev. 104, 1 (1956).

For this simple example we have been able to find the time development of the initial state and explicitly calculate the correlation function $\langle \sigma_{\bar{k}}^+ \sigma_{\bar{k}}^- \sigma_{\bar{k}}^+(t)\, \sigma_{\bar{k}}^-(t) \rangle$. For finite N it exhibits an almost periodic structure. For infinite N it decays to zero like $J_0^2(2\omega t)$. For large N the correlation function behaves like $J_0^2(2\omega t)$ for times less than $\frac{N^2}{4\pi\omega}$. This clearly shows the relationship between the decay and the almost-periodic structures of C(t) for discrete energy levels. We see that this behavior is associated with two properties of the system:

1) The large number of closely spaced energy levels for large N.

2) The large number of energy eigenstates needed in the expansion of the initial state.

We can further see from this example that the dynamic cause of irreversible behavior is related to a dephasing of matrix elements produced by the unitary time transformation.

Correlation Functions for Systems with Continuous Energy Spectra

We shall consider a system having a continuous energy spectrum. Let the eigenstates of the system be represented by $|E\alpha\rangle$.

$$\mathcal{H}\,|E\alpha\rangle = |E\alpha\rangle\, E \tag{6-37}$$

Where α represents the set of additional quantum numbers necessary to specify an eigenstate of the system. For convenience, we shall assume that these additional quantum numbers also have a continuous spectrum. The normalization is then

$$\langle E'\alpha' | E\alpha \rangle = \delta(E'-E)\,\delta(\alpha'-\alpha) \qquad (6\text{-}38)$$

where $\delta(\alpha'-\alpha)$ represents a product of Dirac δ-functions for the set of quantum numbers contained in α. In the case where some of the α quantum numbers are discrete (e.g. polarization or spin) we must replace the Dirac δ-function by a Kronecker δ and the corresponding integration by a summation. Only the continuous nature of the energy spectrum is germane to the following discussion.

Writing the autocorrelation function of F in this representation we find

$$C(t) = \frac{\langle F F(t) \rangle}{\langle F^2 \rangle}$$

$$\langle F F(t) \rangle = \int d\alpha\, d\alpha'\, dE\, dE'\, \langle E\alpha | \rho | E\alpha \rangle\, |\langle E\alpha | F | E'\alpha' \rangle|^2\, e^{i(E'-E)t} \qquad (\hbar = 1)$$

Taking the Fourier transform of C(t) we obtain

$$G(\omega) = \int_{-\infty}^{\infty} dt\, C(t)\, e^{-i\omega t}$$

$$G(\omega) = \frac{2\pi}{\langle F^2 \rangle} \int d\alpha\, d\alpha'\, dE\, dE'\, \langle E\alpha | \rho | E\alpha \rangle\, |\langle E\alpha | F | E'\alpha' \rangle|^2\, \delta(E'-E-\omega) \qquad (6\text{-}39)$$

In the case where F(t) is a random variable, G(ω) is called the "special density" or "power spectrum" of the correlation function C(t); and Eq. (6-39) is then the Wiener-Khintchine theorem. In the present case F(t) is determined by the Heisenberg equations of motion. $G(\omega)$ represents the normalized transition probability density for the system to absorb an energy quantum $\hbar\omega$ in a transition produced by F. This is clear from the following considerations.

$$2\pi \,|\langle E'\alpha' | F | E\alpha\rangle|^2 \,\delta(E' - E - \omega) \qquad (6\text{-}40)$$

is the transition probability (per unit time), first order Born approximation, for the system to go from state $|E\alpha\rangle$ to state $|E'\alpha'\rangle$ with the absorption of energy $\hbar\omega$. Let $\langle E\alpha|\rho|E\alpha\rangle$ be the probability that the system is initially in the state $|E\alpha\rangle$. Then the total transition probability for the absorption of energy $\hbar\omega$ is obtained by averaging Eq. (6-40) over initial states and summing over final states.

$$2\pi \int dE\,dE'\,d\alpha\,d\alpha' \,|\langle E'\alpha'|F|E\alpha\rangle|^2 \,\delta(E' - E - \omega)\,\langle E\alpha|\rho|E\alpha\rangle$$

Normalizing this we obtain $G(\omega)$.

In order to study the long time behavior of the correlation function we consider its Laplace transform.

$$C(s) = \int_0^\infty e^{-st}\, C(t)\,dt$$

$$= \frac{2\pi}{\langle F^2\rangle} \int d\alpha\,d\alpha'\,dE\,dE' \,\frac{\langle E\alpha|\rho|E\alpha\rangle\,|\langle E\alpha|F|E'\alpha'\rangle|^2}{s - i(E' - E)} \qquad (6\text{-}41)$$

In the limit $s \to 0$ we have

$$\lim_{s\to 0} \frac{1}{s-i(E'-E)} = \pi\delta(E-E') - iP\frac{1}{E-E'}$$

Replacing this in Eq. (6-41) we have

$$\lim_{s\to 0} C(s) = \frac{2\pi}{\langle F^2\rangle}\int d\alpha\, d\alpha'\, dE\, dE' \left(\pi\delta(E-E') - iP\frac{1}{E-E'}\right)\langle E\alpha|\rho|E\alpha\rangle\left|\langle E\alpha|F|E'\alpha'\rangle\right|^2$$

Since C(s) is real, the principle value contribution must vanish. This is evidently the case since we may replace

$$\langle E\alpha|\rho|E\alpha\rangle\left|\langle E\alpha|F|E'\alpha'\rangle\right|^2$$

in the integral Eq. (6-39) by the symmetrized form

$$\frac{1}{2}\left(\langle E\alpha|\rho|E\alpha\rangle\left|\langle E\alpha|F|E'\alpha'\rangle\right|^2 + \langle E'\alpha|\rho|E'\alpha'\rangle\left|\langle E'\alpha|F|E\alpha'\rangle\right|^2\right)$$

Therefore we have

$$\lim_{s\to 0} C(s) = \frac{2\pi}{\langle F^2\rangle}\int d\alpha\, d\alpha'\, dE\, dE'\, \langle E\alpha|\rho|E\alpha\rangle\left|\langle E\alpha|F|E'\alpha'\rangle\right|^2\delta(E-E') \quad (6\text{-}42)$$

If $\langle E\alpha|\rho|E\alpha\rangle\left|\langle E\alpha|F|E'\alpha'\rangle\right|^2$ is a well behaved function of E and E' then $\lim_{s\to 0} C(s)$ is finite.[68] This implies that C(t) vanishes as t goes to infinity. This result is in marked contrast to the recurrence property of C(t) for systems with discrete energy spectrums. It is exactly the result one would expect from the previous example of exchange-coupled spins. As an example of a correlation function for a system having a

---

68) This is an important requirement, and essentially characterizes the class of operators and density matrices for which correlations eventually decay. We shall return to it in the last section of this chapter.

continuous energy spectrum we consider $\langle E_x E_x(t,\bar{x})\rangle$ for a thermal equilibrium radiation field.

Electromagnetic Field Correlation Functions

The Hamiltonian of the electromagnetic field in a cubic box of volume $V = L^3$ may be written

$$\mathcal{H} = \sum_{k\lambda} \hbar\omega_k\, a^+_{k\lambda}\, a_{k\lambda} \tag{6-43}$$

where $a_{k\lambda}$ and $a^+_{k\lambda}$ satisfy the commutation relations

$$[a_{k\lambda}, a^+_{k'\lambda'}] = \delta_{kk'}\,\delta_{\lambda\lambda'} \tag{6-44}$$

$$[a_{k\lambda}, a_{k'\lambda'}] = [a^+_{k\lambda}, a^+_{k'\lambda'}] = 0$$

k stands for a vector (k_x,k_y,k_z). Assuming periodic boundary conditions

$$k_x = \frac{2\pi m}{L} \qquad m = 0, \pm 1, \pm 2, \cdots$$

The electromagnetic field operators are given in terms of $a_{k\lambda}$ and $a^+_{k\lambda}$ by the relations[69]

$$\vec{E} = i\sum_{k\lambda}\sqrt{\frac{2\pi\hbar\omega_k}{V}}\,\vec{\epsilon}_{k\lambda}\,(a_{k\lambda}e^{ik\cdot x} - a^+_{k\lambda}e^{-ik\cdot x}) \tag{6-45}$$

$$\vec{H} = i\sum_{k\lambda}\sqrt{\frac{2\pi\hbar c^2}{V\omega_k}}\,(\bar{k}\times\vec{\epsilon}_{k\lambda})\,(a_{k\lambda}e^{ik\cdot x} - a^+_{k\lambda}e^{-ik\cdot x}) \tag{6-46}$$

The time behavior of $a_{k\lambda}$ is determined by

---

69) We are using the wave gauge $\nabla\cdot\vec{A} = 0$, $\phi = 0$.

$$\dot{a}_{k\lambda} = \frac{i}{\hbar}[\mathcal{H}, a_{k\lambda}] = -i\omega a_{k\lambda}$$

$$a_{k\lambda}(t) = e^{-i\omega_k t} a_{k\lambda} \qquad (6\text{-}47)$$

For thermal equilibrium, the state of the radiation field is characterized by the density matrix

$$\rho = \frac{e^{-\beta\mathcal{H}}}{Z(\beta)} \qquad Z(\beta) = \mathrm{Tr}\, e^{-\beta\mathcal{H}}$$

We now calculate the autocorrelation function for the electric field. Since it is a field quantity this correlation function will depend upon space as well as time.

$$C(\vec{x},t) = \frac{\langle E_x\, E_x(\vec{x},t)\rangle}{\langle E_x^2\rangle} = \frac{\mathrm{Tr}\,\rho E_x\, E_x(\vec{x},t)}{\mathrm{Tr}\,\rho E_x^2} \qquad (6\text{-}48)$$

E_x stands for $E_x(0,0)$. Using Eqs. (6-45) and (6-47) the numerator becomes

$$\langle E_x\, E_x(\vec{x},t)\rangle = -\sum_{k\lambda k'\lambda'} \sqrt{\frac{2\pi\hbar\omega_k}{V}}\sqrt{\frac{2\pi\hbar\omega_{k'}}{V}}\,(\vec{\epsilon}_{k\lambda}\cdot\hat{\imath})(\vec{\epsilon}_{k'\lambda'}\cdot\hat{\imath}) \qquad (6\text{-}49)$$

$$\times\, \frac{\mathrm{Tr}\, e^{-\beta\mathcal{H}}(a_{k\lambda} - a_{k\lambda}^{+})(a_{k'\lambda'}\, e^{i(k'\cdot x - \omega_{k'} t)} - a_{k'\lambda'}^{+}\, e^{-i(k'\cdot x - \omega_{k'} t)})}{Z(\beta)}$$

In order for the trace not to vanish we must pair the creation and annihilation operators. The only terms to contribute have the form

$$\mathrm{Tr}\, e^{-\beta\mathcal{H}}\left(a_{k\lambda} a_{k\lambda}^{+}\, e^{-i(k\cdot x - \omega_k t)} + a_{k\lambda}^{+} a_{k\lambda}\, e^{i(k\cdot x - \omega_k t)}\right)$$

Taking the trace we find

$$\frac{\mathrm{Tr}\, e^{-\beta \mathcal{H}} a_{k\lambda} a^+_{k\lambda}}{Z(\beta)} = \bar{n}_k + 1$$

$$\frac{\mathrm{Tr}\, e^{-\beta \mathcal{H}} a^+_{k\lambda} a_{k\lambda}}{Z(\beta)} = \bar{n}_k$$

where

$$\bar{n}_k = \frac{1}{e^{\beta \hbar \omega_k} - 1} \tag{6-50}$$

The expectation value $\langle E_x E_x(\bar{x},t)\rangle$ is therefore

$$\sum_{k\lambda} \left(\frac{2\pi \hbar \omega_k}{V}\right) (\vec{\epsilon}_{k\lambda} \cdot \hat{i})^2 \left[(\bar{n}_k + 1) e^{-i(k\cdot x - \omega_k t)} + \bar{n}_k e^{i(k\cdot x - \omega_k t)} \right] \tag{6-51}$$

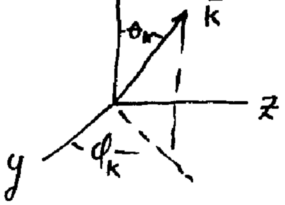
So far we have actually been treating a system which has a discrete spectrum. Now allow the volume V to be infinite and make the replacement

$$\frac{1}{V} \sum_k \longrightarrow \int \frac{d^3k}{(2\pi)^3}$$

We are now dealing with a continuous spectrum. Eq. (6-51) can be written

$$\int \frac{d^3k}{(2\pi)^3} (2\pi \hbar \omega_k) \sum_\lambda (\vec{\epsilon}_{k\lambda} \cdot \hat{i})^2 \left[(\bar{n}_k + 1) e^{-i(k\cdot x - \omega_k t)} + \bar{n}_k e^{i(k\cdot x - \omega_k t)} \right]$$

The angular integration is conveniently performed by choosing the polar axis along the x-direction. The polarization sum becomes



$$\sum_\lambda (\vec{\epsilon}_{k\lambda} \cdot \hat{i})^2 = \sin^2\theta_k$$

$$= \left(\frac{16\pi}{9}\right)^{\frac{1}{2}} Y_{00}(\Omega_k) - \left(\frac{16\pi}{45}\right)^{1/2} Y_{20}(\Omega_k)$$

and the plane wave can be expanded

$$e^{ik\cdot x} = 4\pi \sum_{l=0}^{\infty} i^l j_l(kr) \sum_{m=-l}^{l} Y_{lm}(\Omega_x) Y_{lm}^*(\Omega_k)$$

The Y_{lm} are spherical harmonics. Ω_x refers to the angular coordinates of $\bar{x}$ and Ω_k refers to the angular coordinates of $\bar{k}$. Note that the x-axis has been chosen as the polar axis. j_l are spherical Bessel functions and $r = (x^2 + y^2 + z^2)^{\frac{1}{2}}$. Let H(kr) be defined as the angular integral

$$H(kr) = \frac{3}{8\pi} \int d\Omega_k\, e^{ik\cdot x} \sum_\lambda (\vec{\epsilon}_{k\lambda} \cdot \hat{i})^2 \qquad (6\text{-}52)$$

$$= j_0(kr) - \left(\frac{4\pi}{5}\right)^{1/2} Y_{20}(\Omega_x)\, j_2(kr)$$

The expectation value $\langle E_x E_x(\bar{x},t)\rangle$ is therefore,[70]

$$\langle E_x E_x(\vec{x},t)\rangle = \frac{2\hbar c}{3\pi}\int_0^\infty dk\, k^3 \left[\frac{2\cos kct}{e^{\beta\hbar ck}-1} + e^{ikct}\right] H(kr) \qquad (6\text{-}53)$$

---

[70] Eq. (6-53) illustrates why it is sometimes convenient to define a symmetrized correlation function $\frac{1}{2}\langle E_x E_x(\bar{x},t) + E_x(\bar{x},t)E_x\rangle$. This does not eliminate the singularity in the correlation function, but it does remove its imaginary part.

The second term in Eq. (6-53) corresponds to spontaneous emission. It is singular. Consider an atom going from an excited state $\phi_e(x)$ to its ground state $\phi_0(x)$ by the spontaneous emission of a quantum $\hbar\omega$. The transition probability for this process has the form[71]

$$\int_{-\infty}^{\infty} dt \int d^3x \; |\langle \phi_0(x)|\mu(x)|\phi_e(x)\rangle|^2 \langle E_x E_x(\vec{x},t)\rangle e^{-i\omega t}$$

Carrying out the t integration before the k integration in $\langle E_x E_x(\bar{x},t)\rangle$ we obtain a finite, meaningful result.

For the remainder of this analysis we shall consider only the first (non-singular) part of $\langle E_x E_x(\bar{x},t)\rangle$, Eq. (6-53). We shall denote this by the subscript "1". Let $\gamma = t/\beta\hbar$, $\eta = r/\beta\hbar c$ and $y = \beta\hbar c k$, then

$$\langle E_x E_x(\vec{x},t)\rangle_1 = \frac{4}{3\pi\beta^3\hbar^2c^2}\int_0^\infty dy\, \frac{y^3 \cos\gamma y}{e^y-1} H(\eta y)$$

For $t = \bar{x} = 0$ we have the normalization integral

$$\int_0^\infty dy\, \frac{y^3}{e^y-1} = \frac{\pi^4}{15}$$

The correlation function is therefore,

$$C_1(\vec{x},t) = \frac{\langle E_x E_x(\vec{x},t)\rangle_1}{\langle E_x^2\rangle_1} = \frac{15}{\pi^4}\int_0^\infty dy\, \frac{y^3 \cos\gamma y\, H(\eta y)}{e^y-1} \qquad (6\text{-}54)$$

Inserting $H(\eta y)$ from Eq. (6-52) this becomes

---

[71] We have assumed an interaction of the form $\mu\cdot E$.

$$C_1(\vec{x},t) = \frac{15}{\pi^4}\int_0^\infty dy\, y^3\frac{\cos\gamma y}{e^y-1}\left(j_0(\eta y) - P_2(\cos\theta)\, j_2(\eta y)\right)$$

$$= F_0(r,t) - P_2(\cos\theta)\, F_2(r,t) \tag{6-55}$$

where

$$P_2(\mu) = \frac{1}{2}(3\mu^2-1)$$

$$F_n(r,t) = \frac{15}{\pi^4}\int_0^\infty dy\, y^3\frac{\cos\gamma y}{e^y-1}\, j_n(\eta y) \tag{6-56}$$

The integrals F_0 and F_2 are evaluated in Appendix C. The results are

$$F_0(r,t) = \frac{15}{2\pi\eta}\left(\frac{1}{\pi^3\Gamma_+^3} - \frac{\cosh\pi\Gamma_+}{\sinh^3\pi\Gamma_+} - \frac{1}{\pi^3\Gamma_-^3} + \frac{\cosh\pi\Gamma_-}{\sinh^3\pi\Gamma_-}\right) \tag{6-57}$$

$$F_2(r,t) = \frac{45}{4\pi^3\eta^3}\left(\frac{\cosh\pi\Gamma_+}{\sinh\pi\Gamma_+} - \frac{1}{\pi\Gamma_+} - \frac{\cosh\pi\Gamma_-}{\sinh\pi\Gamma_-} + \frac{1}{\pi\Gamma_-}\right) \tag{6-58}$$

$$-\frac{15}{2\pi\eta}\left(\frac{1}{\pi^3\Gamma_+^3} - \frac{\cosh\pi\Gamma_+}{\sinh^3\pi\Gamma_+} - \frac{1}{\pi^3\Gamma_-^3} + \frac{\cosh\pi\Gamma_-}{\sinh^3\pi\Gamma_-}\right)$$

$$-\frac{45}{4\pi^2\eta^2}\left(\frac{1}{\pi^2\Gamma_+^2} - \frac{1}{\sinh^2\pi\Gamma_+} + \frac{1}{\pi^2\Gamma_-^2} - \frac{1}{\sinh^2\pi\Gamma_-}\right)$$

where

$$\Gamma_\pm = \gamma \pm \eta = \frac{1}{\beta\hbar c}(ct \pm r)$$

The black-body correlation function for vanishing space separation becomes

$$C_1(0,t) = 15\left[\frac{3}{\sinh^4 \pi\gamma} + \frac{2}{\sinh^2 \pi\gamma} - \frac{3}{(\pi\gamma)^4}\right] \tag{6-59}$$

This is plotted in Graph 1. It decays to zero as expected from our previous discussion.[72] The order of magnitude of the decay time is $2\hbar/kT = \omega^{-1}$, where $\hbar\omega = \frac{1}{2}kT$. For $t = 0$, we have the spacial correlation function of black-body radiation.

$$C_1(\vec{x},0) = F_0(r,0) - P_2(\cos\theta)\, F_2(r,0)$$

This is axial symmetric about the x-axis. For $\theta = 0$ it becomes,

$$C_1(X,0) = F_0(X,0) - F_2(X,0)$$

for $\theta = \pi/2$, $x = 0$, $\rho = (y^2 + z^2)^{\frac{1}{2}}$ it is,

$$C_1(\rho,0) = F_0(\rho,0) + \frac{1}{2}F_2(\rho,0)$$

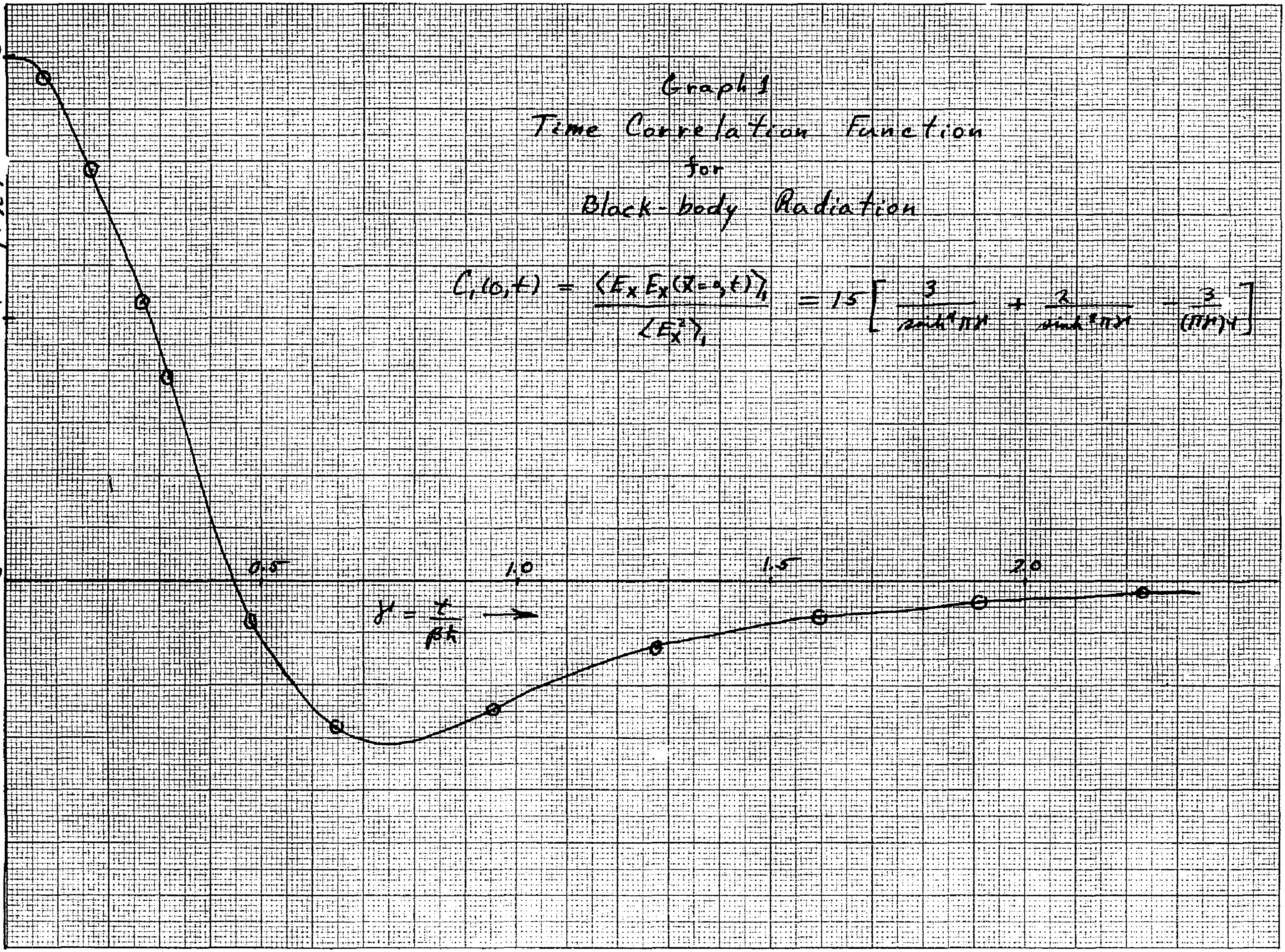
These spatial correlation functions are plotted in Graph 2.

---

72) The Laplace transform of $C_1(0,t)$ has no s^{-1} pole. In fact, it actually vanishes in the limit $s \longrightarrow 0$.

$$C_1(s) = \int_0^\infty dt\, e^{-st} C_1(0,t) = s\,\frac{15}{\pi^4}\int_0^\infty dy\, \frac{y^3}{e^y - 1}\cdot\frac{1}{(y/\beta\hbar)^2 + s^2}$$

$$\lim_{s\to 0} C_1(s) = \lim_{s\to 0} \frac{15}{6}(\beta\hbar)^2 s = 0$$



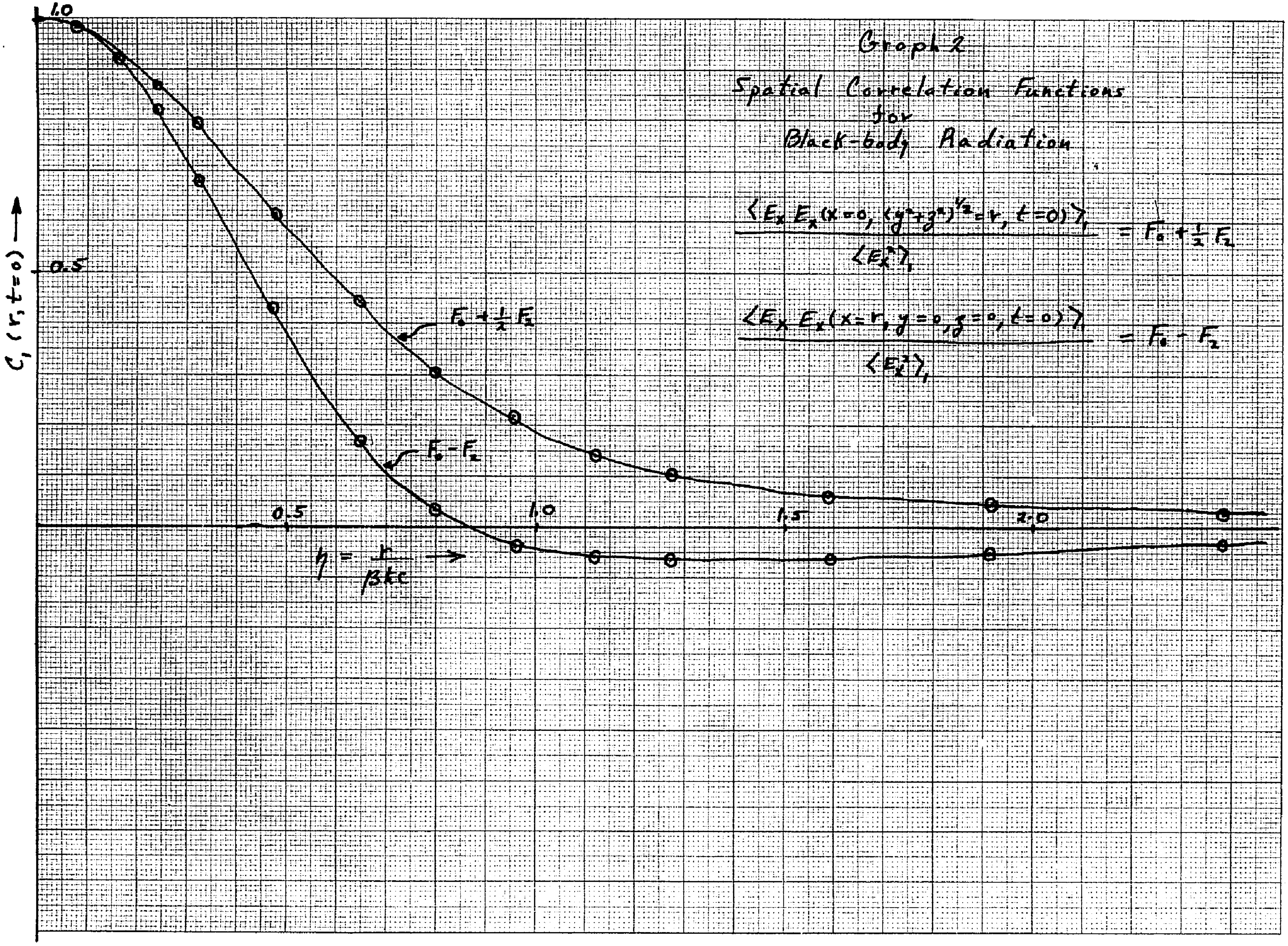
Graph 1

Time Correlation Function

for

Black-body Radiation

$$C_1(0,t) = \frac{\langle E_x E_x(\bar{x}=0,t)\rangle_1}{\langle E_x^2\rangle_1} = 15\left[\frac{3}{\sinh^4 \pi\gamma} + \frac{2}{\sinh^2 \pi\gamma} - \frac{3}{(\pi\gamma)^4}\right]$$



Graph 2

Spatial Correlation Functions
for
Black-body Radiation

$$\frac{\langle E_x E_x(x=0, (y^2+z^2)^{1/2}=r, t=0)\rangle_1}{\langle E_x^2\rangle_1} = F_0 + \frac{1}{2}F_2$$

$$\frac{\langle E_x E_x(x=r, y=0, z=0, t=0)\rangle_1}{\langle E_x^2\rangle_1} = F_0 - F_2$$

Further second order correlation functions for the field variables $\overline{E}$ and $\overline{H}$ can be calculated in the same manner. The major differences occur as a result of the polarization sum. We conclude this section by noting several of these.

$$\langle E_x E_y(\vec{x},t)\rangle \sim \sum_{\lambda} (\vec{\epsilon}_\lambda\cdot\hat{\imath})(\vec{\epsilon}_\lambda\cdot\hat{\jmath}) = -\sin\theta_k\cos\theta_k\sin\phi_k$$

For $\overline{x} = 0$ this expectation value vanishes. We shall therefore normalize this correlation function with respect to $\langle E_x^2\rangle_1$.

$$\frac{\langle E_x E_y(\vec{x},t)\rangle_1}{\langle E_x^2\rangle_1} = \frac{3}{2}\sin\theta\cos\theta\sin\phi\, F_2(r,t)$$

In a similar manner we find

$$\frac{\langle E_x E_z(\vec{x},t)\rangle_1}{\langle E_x^2\rangle_1} = \frac{3}{2}\sin\theta\cos\theta\cos\phi\, F_2(r,t)$$

The polarization sum for $\langle H_x H_x(\overline{x},t)\rangle$ is $\sin^2\theta_k$. This has the same angular dependence as the polarization sum for $\langle E_x E_x(\overline{x},t)\rangle$. Therefore their correlation functions are identical. The polarization sum for $\langle H_x E_x(\overline{x},t)\rangle$ vanishes. The polarization sum for $\langle H_y E_x(\overline{x},t)\rangle$ is

$$\langle H_y E_x(\vec{x},t)\rangle \sim \sum_{\lambda} (\vec{k}\times\vec{\epsilon}_\lambda)\cdot\hat{\jmath}\,(\epsilon_\lambda\cdot\hat{\imath}) = k\sin\phi_k\sin\theta_k$$

and we find the correlation function

$$\frac{\langle H_y E_x(\vec{x},t)\rangle_1}{\langle E_x^2\rangle_1} = \frac{3}{2}\sin\phi\sin\theta\, F_1'(r,t)$$

where

$$F_1'(r,t) = \frac{15}{\pi^4}\int_0^\infty dy \, \frac{y^3 \sin \delta' y}{e^y - 1} \, j_1(\eta y)$$

The Approach to Equilibrium of Macroscopic Observables

From the preceding discussion of correlation functions we have seen that the dynamic cause of their irreversible decay is a dephasing, and consequent destructive interference among the terms of a sum of matrix elements. For a system having a finite number of degrees of freedom the character of this interference will eventually become constructive and the system will return arbitrarily close to its initial state. However, for large systems this Poincaré recurrence time is many orders of magnitude larger than any laboratory time of interest. Formally it may be mathematically convenient to consider the asymptotic time behavior. If this is the case we have seen that one must take the limit $N \rightarrow \infty$ before taking the limit $t \rightarrow \infty$. In this way the almost-periodic structure of the matrix element is eliminated. We shall now use these notions to investigate the irreversible tendency of macroscopic systems to approach equilibrium. In particular we shall be interested in the properties of $\mathcal{H}$, ρ and F which are necessary in order for $\mathrm{Tr}\,\rho F(t)$ to approach equilibrium.

Let the time development of a system be determined by $\mathcal{H}$. We initially consider the case of a discrete spectrum

$$\mathcal{H}|n\rangle = \omega_n |n\rangle \qquad (6\text{-}60)$$
$$\langle n'|n\rangle = \delta_{n'n}$$

If the initial state of the system is ψ_0, the time evolution of the dynamic variable F is

$$\langle\psi_0|F(t)|\psi_0\rangle = \langle\psi_0|e^{i\mathcal{H}t} F e^{-i\mathcal{H}t}|\psi_0\rangle \qquad (\hbar = 1)$$

$$= \sum_{nm} \langle\psi_0|e^{i\mathcal{H}t}|n\rangle\langle n|F|m\rangle\langle m|e^{-i\mathcal{H}t}|\psi_0\rangle$$

$$= \sum_{nm} \langle\psi_0|n\rangle\langle n|F|m\rangle\langle m|\psi_0\rangle e^{i\omega_{nm}t} \qquad (6\text{-}61)$$

where $\omega_{nm} = \omega_n - \omega_m$. At t = 0 all the terms in Eq. (6-61) will add in a constructive manner. As t increases we are interested in the interference which causes $\langle\psi_0|F(t)|\psi_0\rangle$ to change from its initial value to a final equilibrium value. The long time character of this interference depends upon the distribution of energy level spacings ω_{nm} which contribute to the sum (6-61). For initial states ψ_0 and operators F such that (6-61) consists of only a small number of terms, the interference effects will give rise to oscillations on a laboratory time scale. For the alternate case in which the sum (6-61) consists of a large number of exponential terms the interference effects are observed as an irreversible change on a laboratory time scale. Necessary conditions for this latter effect are: 1) the expansion of ψ_0 in terms of the energy eigenstates $|n\rangle$ contains a large number of terms;[73] 2) F has a large number of off diagonal matrix elements in the energy representation;[74] 3) the energy levels are closely spaced.

In order to obtain a more quantitative criterion for irreversible change, consider the Laplace transform of Eq. (6-61).

$$F(s) = \int_0^\infty dt\, e^{-st} \langle \psi_0 | F(t) | \psi_0 \rangle = \sum_{nm} \frac{\langle \psi_0 | n \rangle \langle n | F | m \rangle \langle m | \psi_0 \rangle}{s - i\omega_{nm}} \qquad (6\text{-}62)$$

F(s) has a series of poles located along the imaginary s-axis. $\langle \psi_0 | F(t) | \psi_0 \rangle$ is an almost-periodic function. As previously noted, the physically relevant, asymptotic, time behavior is most easily found by going to the limiting case of a continuous spectrum. In this case the poles become continuously distributed. The time behavior is then determined by the density of poles. We illustrate this by some examples.

Consider the case in which there are 2N+1 poles uniformly distributed along the imaginary s-axis between ± i.

---

73) In actual practice the initial state of the system is often represented by a density matrix. This implies an additional summation and the possibility for an increase in the number of exponential terms contributing to the sum Eq. (6-61).

74) Most macroscopic observables consist of the sum of a large number of one or two particle operators. This is the reason they satisfy condition (2).

$$F_N(s) = \frac{1}{2N+1} \sum_{m=-N}^{N} \frac{1}{s - i\frac{m}{N}} \tag{6-63}$$

The inverse Laplace transform of this is

$$f_N(t) = \frac{1}{2N+1} \sum_{m=-N}^{N} e^{i\frac{m}{N}t}$$

In the limit of large N we find

$$f(t) \equiv \lim_{N\to\infty} f_N(t) = \frac{1}{2}\int_{-1}^{1} dx\, e^{ixt} = \frac{\sin t}{t}$$

and f(t) decays as t^{-1}.

Next consider the distribution generated by projecting onto the imaginary axis a uniform distribution of N points on the unit circle.

$$F_N(s) = \frac{1}{N} \sum_{m=1}^{N} \frac{1}{s - i\cos\frac{2\pi m}{N}} \tag{6-64}$$

The inverse Laplace transform of this is

$$f_N(t) = \frac{1}{N} \sum_{m=1}^{N} e^{i(\cos\frac{2\pi m}{N})t}$$

In the limit of large N we find

$$f(t) \equiv \lim_{N\to\infty} f_N(t) = \frac{1}{2\pi}\int_{0}^{2\pi} d\theta\, e^{i(\cos\theta)t} = J_0(t)$$

where $J_0(t)$ is the zero order Bessel function.

$$\lim_{t\to\infty} J_0(t) \sim \left(\frac{2}{\pi t}\right)^{1/2} \cos\left(t - \frac{\pi}{4}\right)$$

Therefore f(t) decays like $t^{-\frac{1}{2}}$. The difference between this case and the previous one can be seen by examining the distribution of poles. Let $s = x + iy$, then for the first case the distribution of poles is

$$m(y) = \begin{cases} \frac{1}{2} & |y| \leq 1 \\ 0 & |y| > 1 \end{cases}$$

For the second case it is

$$m(y) = \begin{cases} \frac{1}{2\pi} \frac{1}{(1-y^2)^{1/2}} & |y| \leq 1 \\ 0 & |y| > 1 \end{cases}$$

The slower time decay associated with the latter distribution arises from the singularities at $y = \pm 1$.

Finally, consider the distribution of poles

$$m(y) = \frac{e^{-\frac{(y-1)^2}{a^2}}}{2\pi^{1/2} a} + \frac{e^{-\frac{(y+1)^2}{a^2}}}{2\pi^{1/2} a} \qquad -\infty < y < \infty \tag{6-65}$$

The inverse Laplace transform is

$$f(t) = \int_{-\infty}^{\infty} dy\, m(y)\, e^{iyt} = e^{-\left(\frac{ta}{2}\right)^2} \cos t$$

The decay time of f(t) varies as a^{-1}. In the limit $a \to 0$, f(t) does not decay. In this limit the distribution of poles becomes

$$\lim_{a \to 0} m(y) = \frac{1}{2}\left(\delta(y+1) + \delta(y-1)\right) \tag{6-66}$$

These examples illustrate the importance of the singularities which occur in the density of poles distribution function. It is the δ-function singularities which are responsible for the asymptotic time behavior of an expectation value. The effects of singularities which are weaker than a δ-function vanish in the asymptotic time limit. We shall now construct the density of poles distribution function which determines the time development of the expectation value of a dynamic variable F.

We assume the energy spectrum of the system is continuous. Let the initial state be ψ_0. Using the basis, Eq. (6-37), we have

$$\langle\psi_0|F(t)|\psi_0\rangle = \int dE\,dE'\,d\alpha\,d\alpha'\; e^{i(E-E')t} \langle\psi_0|E\alpha\rangle\langle E\alpha|F|E'\alpha'\rangle\langle E'\alpha'|\psi_0\rangle \quad (6\text{-}67)$$

The Laplace transform of Eq. (6-67) is

$$F(s) = \int_0^\infty dt\; e^{-st}\langle\psi_0|F(t)|\psi_0\rangle$$

$$F(s) = \int dE\,dE'\,d\alpha\,d\alpha'\; \frac{\langle\psi_0|E\alpha\rangle\langle E\alpha|F|E'\alpha'\rangle\langle E'\alpha'|\psi_0\rangle}{s-i(E-E')}$$

$$F(s) = \int dE''\,\frac{M(E'')}{s-iE''} \quad ; E''=E-E'$$

The density of poles distribution function is $M(E'') = \int dE'\,F(E''+E',E')$ where,

$$F(E,E') = \int d\alpha\,d\alpha'\,\langle\psi_0|E\alpha\rangle\langle E\alpha|F|E'\alpha'\rangle\langle E'\alpha'|\psi_0\rangle \quad (6\text{-}68)$$

From the discussion examples it appears that one can implicitly state requirements (1), (2) and (3) (p. 107) in terms of the behavior of F(E,E').[75] If F(E,E') has no singularities as strong as a δ-function, then $\langle F(t)\rangle$ asymptotically decays to zero. If F(E,E') has the form

$$F(E,E') = F_1(E)\,\delta(E-E') + F_2(E,E') \tag{6-69}$$

where $F_2(E,E')$ has no singularities as strong as a δ-function, then $\langle F(t)\rangle$ asymptotically approaches the equilibrium value

$$\langle F(t)\rangle_{eq.} = \int dE\, F_1(E) \tag{6-70}$$

If F(E,E') has a singularity of the type

$$\frac{1}{2}\, F^{\omega}(E)\,\big(\delta(E-E'+\omega)+\delta(E-E'-\omega)\big) \tag{6-71}$$

then $\langle F(t)\rangle$ will have a term

$$\cos\omega t \cdot \int dE\, F^{\omega}(E) \tag{6-72}$$

Unless

$$\int dE\, F^{\omega}(E) = 0 \tag{6-73}$$

the long time behavior of $\langle F(t)\rangle$ will be oscillatory.

In general we expect the density of poles function for a macroscopic system to have the form shown in Eq. (6-69). The δ- singularity arises from the group of matrix elements

---

75) See H. Longuet-Higgins and S. Golden, Journ. of Chem. Phys. 33, 1479 (1960).

having E equal to E'. It is just this group of matrix elements which maintain their phase relationship in time. The other terms in the series, Eq. (6-67), become dephased due to the time dependent exponential phase factors exp(i(E-E')t).

APPENDIX A

Three Fundamental Theorems of Information Theory

In this appendix we will state three theorems of information theory originally given by C. E. Shannon.76) These theorems provide a background for the information theory concepts used in this thesis. The first theorem is concerned with the uniqueness of the entropy functional as a measure of the uncertainty associated with a probability distribution. The next two theorems are related to the existence of, and the probabilities associated with, a certain class of sequences generated by a Markov process.

Shannon originally formulated information theory to provide a mathematical theory of communication. The basic problems of communication theory are concerned with the transfer of information by means of messages. Khinchin77) has since given a discussion of Shannon's work in which some of the practical details associated with communication problems have been avoided. The following discussion will draw heavily from both these sources.

Consider a system which has only two possible states s_1 and s_2. Let p_1 and p_2 represent the probability that the system is in the states s_1 and s_2 respectively. Following Khinchin we shall call this a finite scheme and denote it by

---

76) C. E. Shannon, op. cit.

77) A. I. Khinchin, Mathematical Foundations of Information Theory (Dover Publications, Inc., New York, 1957).

$$\begin{pmatrix} s_1 & s_2 \\ p_1 & p_2 \end{pmatrix}$$

One characteristic of a finite scheme is the uncertainty it represents. To illustrate this consider the two schemes

$$\begin{pmatrix} s_1 & s_2 \\ 0.5 & 0.5 \end{pmatrix} \begin{pmatrix} s_1 & s_2 \\ .99 & .01 \end{pmatrix}$$

Given the first scheme one feels uncertain as to the state in which the system will be found. Given the second scheme one feels quite certain that the system will be found in state s_1. This same property can be described by saying that the result of a measurement determining the state of the system conveys more information in the first case than in the second.

The first theorem of interest is concerned with prescribing a unique measure for the uncertainty (or information obtained by a measurement) associated with the finite scheme S.

$$S = \begin{pmatrix} s_1 & s_2 & \cdots & s_N \\ p_1 & p_2 & \cdots & p_N \end{pmatrix} \qquad \text{(A-1)}$$

If such a measure, $H(p_1, p_2, \ldots p_N)$, (called the entropy) exists, it should satisfy certain consistency conditions. The most important of these is the composition law which results from the requirement that the entropies of equivalent finite schemes should be equal. Instead of giving the finite scheme S directly, one might give a finite scheme S_o consisting of a set of composite states (e.g. The first composite state consists of the states A_1, A_2 and A_3 and the probability p_1' of this composite state is the sum of the probabilities of the system being in the individual states A_1, A_2 or A_3. The second composite state consists of the states A_4 and A_5 etc.) Let the entropy associated with S_0 be $H(S_0)$. Now suppose in addition to S_0 one has all the conditional schemes S_i containing the conditional probabilities for the individual states making up the $i^{\underline{th}}$ composite state given that the $i^{\underline{th}}$ composite state has occurred. Let the entropy of S_i be $H(S_i)$. The mutually dependent schemes S_0, S_1, ... contain the same information as the original scheme S. They are therefore equivalent schemes and we obtain the composition law

$$H(S) = H(S_0) + p_1'H(S_1) + p_2'H(S_2) + \dots$$

where p_i' represents the probability associated with the $i^{\underline{th}}$ composite state as given by S_0. The weight factors p_i' occur because the additional entropy associated with the $i^{\underline{th}}$ composite state scheme is encountered only with probability p_i'.

Another consistency condition necessary to provide agreement with common sense is: if all p_i are equal, $p_i = \frac{1}{N}$, then H is a monotonic increasing function of N. The final requirement on H is that it be a continuous function of the p_i. We are now in a position to state the first theorem.[78)]

Theorem I The Uniqueness Theorem: The only H satisfying the three consistency conditions has the form

---

78) Theorems I, II and III of this Appendix correspond to Shannon's theorems 2, 4, and 3 respectively.

$$H = -k \sum_{i=1}^{N} p_i \log p_i$$

where k is a positive constant.

In the following discussion we shall set k equal to one and use natural logarithms.

Shannon's first paper is concerned with messages generated by a stationary, finite, Markov chain. A stationary, finite, Markov chain is characterized by a set of possible states s_1, s_2, ...s_N and a transition probability matrix p_{ik}. Where p_{ik} denotes the conditional probability that s_k will be the next state in the chain if the present state is s_i. Let the probability of the state s_i be denoted by P_i. If the system is in state s_i, then the transition probabilities p_{ik} form a finite scheme

$$\begin{pmatrix} s_1 & s_2 & \cdots & s_N \\ p_{i1} & p_{i2} & \cdots & p_{iN} \end{pmatrix}$$

The entropy of this scheme is

$$H_i = -\sum_{k=1}^{N} p_{ik} \log p_{ik}$$

The average entropy per state in the chain is

$$H = \sum_{i=1}^{N} P_i H_i \qquad \text{(A-2)}$$

The applications of Theorem I in communication theory are based upon the relationship between H and a special class of high probability sequences. Sequences are formed by taking consecutive states of a Markov chain. For a Markov chain having N different states there are N^r different sequences containing r states. These sequences have the form

$$s_{k_1} s_{k_2} \cdots s_{k_r} \qquad \text{(A-3)}$$

where k_i takes on the values 1 to N. The probability associated with the sequence (A-3) is

$$P(r) = P_{k_1} P_{k_1 k_2} \cdots P_{k_{r-1} k_r} \qquad \text{(A-4)}$$

In Shannon's work the existence of a special class of sequences is proved for the case in which the Markov process is ergodic.79),80) Ergodic means that the Markov process obeys the law of large numbers. That is, the fraction of times a given state s_i occurs in a sequence of length r will differ from P_i by an arbitrarily small amount with a probability arbitrarily close to unity for sufficiently large r.

Arrange the N^r different sequences of the form (A-3) in order of decreasing probability (A-4). Select in order,

---

79) A process is ergodic if it is possible to make a transition from any state to any other state in a finite number of steps.

80) R. Nelson (op.cit.) has investigated the existence of this special class of sequences for more general processes. He has also produced a counter example of a process for which the special class does not exist.

starting from the most probable, a set of sequences. Let $N_r(\lambda)$ denote the smallest number of sequences such that their total probability P satisfies $P \geq \lambda$ $(0 < \lambda < 1)$. These sequences form the "high probability" class. The "high probability" class is sharply defined in the following sense.

Theorem II: For an ergodic Markov chain

$$\lim_{r \to \infty} \frac{\log N_r(\lambda)}{r} = H \quad (\lambda \neq 1,0)$$

where H is independent of λ, and is defined by Eq. (A-2).

Thus in the limit of long sequences (large r) the number of sequences in the "high probability" class is approximately e^{rH} independent of how we choose to define "high probability" (e.g. whether we choose λ = .01 or .99).

A related theorem deals with the probabilities associated with this selected class of "high probability" sequences.

Theorem III: Given any $\epsilon > 0$ and $\delta > 0$, we can find an r_o such that the sequences of any length $r \geq r_o$ fall into two classes:

1. A set whose total probability is less than ϵ (the "low probability" class).
2. The remainder (the "high probability" class), all of whose members have probabilities $P(r)$, Eq. (A-3), satisfying the inequality

$$\left| \frac{\log P(r)^{-1}}{r} - H \right| < \delta$$

Thus the probability associated with almost all high probability sequences is very close to e^{-Hr}, when r is large.

These two theorems imply that for most purposes it is possible to treat the long sequences of r states as though there were just 2^{Hr} of them, each with a probability 2^{-Hr}.

APPENDIX B

The Energy Density of States Function for a Spin System

Let $N(E)\Delta E$ be the number of global quantum states lying in the energy interval $(E, E + \Delta E)$. $N(E)$ is called the energy density of states function. We wish to calculate $N(E)$ for a system of N spins located in a rigid lattice. First consider the case of non-interacting spins. If the lattice is located in an external field H, the energy of the spin system is

$$E = -\mu H \sum_{m=-I}^{I} N_m m \qquad \text{(B-1)}$$

N_m represents the number of spins having a z-component of angular momentum $\hbar m$. m takes on the 2I+1 values -I to +I. The magnetic moment of a spin is μI. For a system composed of N spins

$$\sum_{m=-I}^{I} N_m = N \qquad \text{(B-2)}$$

The energy density of states function for this system of non-interacting spins is

$$N(E) = \frac{1}{\mu H} \sum_{(N_m)}{}' N! \left[\prod_{m=-I}^{I} N_m ! \right]^{-1} \qquad \text{(B-3)}$$

The sum is over the sets (N_m) which satisfy Eqs. (B-1) and (B-2).

The sum, Eq. (B-3), can be evaluated using the Darwin-

Fowler method.[81] Define the generating function

$$W(x,y) = \sum_{(N_m)} N! \left[\prod_{m=-I}^{I} N_m! \right]^{-1} x^{\left(\sum_{m=-I}^{I} N_m\right)} y^{\left(-\sum_{m=-I}^{I} m N_m\right)} \tag{B-4}$$

where the summation is unrestricted. Now imagine that the generating function W(x,y) is expended in a power series

$$W(x,y) = \sum_{\nu,\lambda} W_{\nu,\lambda}\, x^{\nu} y^{\lambda}$$

The sum in Eq. (B-3) is given by the coefficient for which $\nu = N$ and $\lambda = \frac{E}{\mu H}$. Using the theory of complex variables we have

$$W_{N,\frac{E}{\mu H}} = \left(\frac{1}{2\pi i}\right)^2 \oint dy \oint dx \; \frac{W(x,y)}{x^{(N+1)}\, y^{\left(\frac{E}{\mu H}+1\right)}} \tag{B-5}$$

where the contours are taken counterclockwise about the origin. It is assumed they lie within the circles of convergence of W(x,y).

Summing Eq. (B-4) we obtain

$$W(x,y) = N!\, e^{x f(y)} \tag{B-6}$$

where

$$f(y) = \sum_{m=-I}^{I} y^{-m} = \frac{\sinh\left(\frac{2I+1}{2I}\eta\right)}{\sinh\left(\frac{1}{2I}\eta\right)} \tag{B-7}$$

$$\eta = I \ln y$$

---

[81] R. H. Fowler, Statistical Mechanics (Cambridge Univ. Press, London, 1936).

The x-integral in Eq. (B-5) can be done by inspection and we obtain

$$W_{N,\frac{E}{\mu H}} = \frac{1}{2\pi i} \oint dy \, \frac{f^N(y)}{y^{\left(\frac{E}{\mu H}+1\right)}} \tag{B-8}$$

This integral is evaluated by the method of steepest descent. Define g(y) by the relation,

$$e^{\frac{E}{\mu H} g(y)} = \frac{f^N(y)}{y^{\left(\frac{E}{\mu H}+1\right)}} \tag{B-9}$$

Then

$$W_{N,\frac{E}{\mu H}} \doteq \frac{e^{\frac{E}{\mu H} g(y_0)}}{\left(2\pi \frac{E}{\mu H} g''(y_0)\right)^{1/2}} \tag{B-10}$$

where y_0 is determined by the condition

$$g'(y_0) = 0 \tag{B-11}$$

In terms of f Eq. (B-11) becomes

$$y_0 \frac{f'(y_0)}{f(y_0)} = \frac{E}{N\mu H} \tag{B-12}$$

Using Eq. (B-7) this becomes

$$I\, B_I(\eta_0) = \frac{E}{N\mu H} \tag{B-13}$$

where $B_I(\eta_0)$ is the Brillouin function

$$B_I(\eta_0) = \frac{2I+1}{2I} \coth\left(\frac{2I+1}{2I}\eta_0\right) - \frac{1}{2I}\coth\frac{1}{2I}\eta_0 \tag{B-14}$$

$$\eta_0 = I \ln y_0$$

For small η_0, $B_I(\eta_0)$ may be expanded

$$B_I(\eta_0) \doteq \frac{1}{3} I(I+1) \ln y_0 \tag{B-15}$$

Substituting (B-15) into (B-13) we obtain $\ln y_0$.

$$\ln y_0 \doteq \frac{3E}{N\mu H I(I+1)} \tag{B-16}$$

For the magnetic field strengths of interest $M \ll N\mu I$. Therefore $\ln y_0 \ll 1$ and the expansion (B-15) is an excellent approximation.

The calculation is completed by evaluating $g(y_0)$ and $g''(y_0)$. We find

$$\frac{E}{\mu H} g(y_0) = N \ln(2I+1) - \frac{3E^2}{2N\mu^2 H^2 I(I+1)} - \ln y_0$$

$$g''(y_0) = \frac{N\mu H I(I+1)}{3 E y_0^2}$$

Substituting these expressions into Eq. (B-10) and dividing by μH we obtain the energy density of states function

$$N(E) = \frac{(2I+1)^N \exp\left[-\frac{3E^2}{2N\mu^2H^2I(I+1)}\right]}{\left(\frac{2\pi N\mu^2H^2I(I+1)}{3}\right)^{1/2}} \qquad \text{(B-17)}$$

We note that N(E) is correctly normalized

$$\int_{-\infty}^{\infty} N(E)\,dE = (2I+1)^N$$

since we have assumed that the individual spins can be distinguished by their lattice sites.

We cannot use this procedure to determine N(E) when the spins are coupled by a dipole-dipole interaction. For this case it is convenient to define the energy density of states in terms of the partition function.

$$\int_{-\infty}^{\infty} N(E)\,e^{-\beta E}\,dE = Tr\, e^{-\beta\mathcal{H}} \qquad \text{(B-18)}$$

where

$$\mathcal{H} = -HM + \mathcal{D} \qquad \text{(B-19)}$$

$$M = \mu\sum_{i=1}^{N} I_{z_i} \qquad \mathcal{D} = \sum_{i>j}\sum_{\mu\nu} D_{\mu\nu}(ij)\, I_i^{\mu} I_j^{\nu}$$

The partition function can be formally expanded in powers of β. This corresponds to the "high temperature" approximation and will lead to physical results which are valid when the energy per spin is small compared to kT. Noting that Tr$\mathcal{H}$ = 0, we obtain in lowest order

$$\int_{-\infty}^{\infty} N(E)\, e^{-\beta E} dE \doteq \mathrm{Tr}\,1 \left(1 + \tfrac{1}{2} \beta^2 \frac{\mathrm{Tr}\,\mathcal{H}^2}{\mathrm{Tr}\,1} \right) \tag{B-20}$$

Since TrMD = 0,

$$\frac{\mathrm{Tr}\,\mathcal{H}^2}{\mathrm{Tr}\,1} = k C \left(H^2 + H_L^2 \right) \tag{B-21}$$

where C is Curie's constant

$$C = \frac{1}{k} \frac{\mathrm{Tr}\,M^2}{\mathrm{Tr}\,1} = \frac{N \mu^2 I\,(I+1)}{3k} \tag{B-22}$$

and H_L is the local dipole-dipole field defined by

$$H_L^2 = \frac{\mathrm{Tr}\,D^2}{\mathrm{Tr}\,M^2} \tag{B-23}$$

We solve Eq. (B-20) by the <u>Ansatz</u>

$$N(E) = \frac{N_0}{\pi^{1/2} \Delta^{1/2}} e^{-\frac{E^2}{\Delta}} \tag{B-24}$$

The integral becomes

$$\int_{-\infty}^{\infty} N(E)\, dE = N_0\, e^{\frac{\beta^2 \Delta}{4}} \tag{B-25}$$

Expand this in powers of β, and equate coefficients with the right-hand side of Eq. (B-20). We find to order β^2

$$N_0 = \mathrm{Tr}\,1$$

$$\Delta = 2kC\,(H^2 + H_L^2) \qquad \text{(B-26)}$$

We conclude by noting several properties of N(E), Eqs. (B-24) and (B-26). Using Gibbs' second analogy definition of temperature

$$\frac{1}{kT} = \frac{d}{dE}\ln N(E) = -\,\frac{E}{kC\,(H^2+H_L^2)}$$

Therefore, the energy is given in terms of T and H by,

$$E = -\,\frac{C}{T}\,(H^2 + H_L^2)$$

This is just the energy relationship obtained from the density matrix

$$\rho = \frac{e^{-\beta\mathcal{H}}}{Z(\beta)} \qquad Z(\beta) = \mathrm{Tr}\,e^{-\beta\mathcal{H}} \qquad \beta = \frac{1}{kT}$$

using the "high temperature" approximation. In the limiting case $H_L \to 0$, N(E) for the interacting system reduces to the energy density of states function for the non-interacting system, Eq. (B-17). In the approximation used, the spin-spin interaction merely broadens the distribution of energy levels by a factor $\left(1 + \frac{H_L^2}{H^2}\right)^{\frac{1}{2}}$. It does not affect its Gaussian shape.

APPENDIX C

Black-Body Correlation Function Integrals

The integrals occurring in the electromagnetic field correlation functions can be obtained by differentiating I(a).

$$I(a) = \int_0^\infty dy \, \frac{\sin ay}{e^y - 1} = \frac{\pi}{2}\left(\coth \pi a - \frac{1}{\pi a}\right)$$

This integral is evaluated by expanding the denominator

$$\frac{1}{e^y - 1} = \sum_{m=1}^{\infty} e^{-my}$$

carrying out the y integration

$$I(a) = \sum_{m=1}^{\infty} \frac{a}{a^2 + m^2}$$

and summing the series

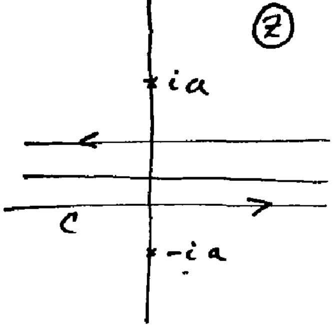
$$I(a) = \sum_{m=1}^{\infty} \frac{a}{a^2+m^2} = \frac{\pi}{2}\left(\coth \pi a - \frac{1}{\pi a}\right)$$

The series is summed by converting it to an integral in the complex plane.

$$\sum_{m=1}^{\infty} \frac{a}{a^2+m^2} = \frac{1}{2}\left[\sum_{m=-\infty}^{\infty} \frac{a}{a^2+m^2} - \frac{1}{a}\right]$$

②

$$\sum_{m=-\infty}^{\infty} \frac{a}{a^2+m^2} = \frac{1}{2\pi i} \oint_C dz \, \frac{\pi a}{a^2+z^2} \cot z\pi$$



The contour C can be continuously deformed into circles around the poles ±ia.

$$\frac{1}{2\pi i} \oint_C dz \, \frac{\pi a \cot z\pi}{a^2+z^2} = -\lim_{z \to ia} \frac{\pi a \cot \pi z}{z+ia} - \lim_{z \to -ia} \frac{\pi a \cot \pi z}{z-ia}$$

$$= -\frac{\pi}{i} \cot i\pi a$$

$$= \pi \coth \pi a$$

The integrals appearing in $\langle E_x E_x(\bar{x},t) \rangle$ are:

$$F_0 = \frac{15}{\pi^4} \int_0^\infty dy \, \frac{y^3 \cos \gamma y}{e^y - 1} \, j_0(\eta y) = \frac{15}{2\pi^4 \eta} \left(-I''(\gamma+\eta) + I''(\gamma-\eta) \right)$$

$$= \frac{15}{2\pi \eta} \left(-\frac{\cosh \pi \Gamma_+}{\sinh^3 \pi \Gamma_+} + \frac{1}{\pi^3 \Gamma_+^3} + \frac{\cosh \pi \Gamma_-}{\sinh^3 \pi \Gamma_-} - \frac{1}{\pi^3 \Gamma_-^3} \right)$$

$$F_2 = \frac{15}{\pi^4} \int_0^\infty dy \, \frac{y^3 \cos \gamma y}{e^y - 1} \, j_2(\eta y) = \frac{45}{2\pi^4 \eta^3} \left(I(\gamma+\eta) - I(\gamma-\eta) \right)$$

$$- \frac{15}{2\pi^4 \eta} \left(-I''(\gamma+\eta) + I''(\gamma-\eta) \right) - \frac{45}{2\pi^4 \eta^2} \left(I'(\gamma+\eta) + I'(\gamma-\eta) \right)$$

$$F_2 = \frac{45}{4\pi^3\eta^3}\left(\frac{\cosh \pi\Gamma_+}{\sinh \pi\Gamma_+} - \frac{1}{\pi\Gamma_+} - \frac{\cosh \pi\Gamma_-}{\sinh \pi\Gamma_-} + \frac{1}{\pi\Gamma_-}\right).$$

$$- \frac{15}{2\pi\eta}\left(-\frac{\cosh \pi\Gamma_+}{\sinh^3 \pi\Gamma_+} + \frac{1}{\pi^3\Gamma_+^3} + \frac{\cosh \pi\Gamma_-}{\sinh^3 \pi\Gamma_-} - \frac{1}{\pi^3\Gamma_-^3}\right)$$

$$- \frac{45}{4\pi^2\eta^2}\left(-\frac{1}{\sinh^2 \pi\Gamma_+} + \frac{1}{\pi^2\Gamma_+^2} - \frac{1}{\sinh^2 \pi\Gamma_-} + \frac{1}{\pi^2\Gamma_-^2}\right)$$

where $\Gamma_\pm = \gamma \pm \eta$ and $I'(a) = \frac{d}{da} I(a)$. For the case $r = 0$ only the first integral contributes. It becomes

$$\frac{15}{\pi^4}\int_0^\infty dy\, \frac{y^3 \cos \gamma y}{e^y - 1} = 15\left(\frac{3}{\sinh^4 \pi\gamma} + \frac{2}{\sinh^2 \pi\gamma} - \frac{3}{\pi^4\gamma^4}\right)$$