REVIEWS OF MODERN PHYSICS VOLUME 29, NUMBER 1 JANUARY, 1957

Description of States in Quantum Mechanics by Density Matrix and Operator Techniques

U. FANO

National Bureau of Standards, Washington, D. C.

Contents

1. INTRODUCTION

To describe a phenomenon correctly and conveniently, one wants to represent the states of physical systems by parameters that are logically consistent and also have a familiar, operational significance. Difficulties in achieving this goal have come up, for example, in the study of atomic phenomena involving polarization, spin orientations, and angular correlations. In these phenomena, and in many others, the experimental procedure does not usually analyze all relevant variables to the maximum extent consistent with quantum mechanics. We refer to states of "less than maximum information," to indicate for example that we know less about the spin orientation of the protons in a drop of water than we would if each proton were in a state characterized by a magnetic quantum number.

The difficulty in familiarizing oneself with states of less-than-maximum information leads, for example, to ask: "Given a beam of unpolarized electrons, should one think of each electron as having a definite spin orientation?" Questions of this kind are semantically slippery, of course, and should be argued in terms of some specific experimental arrangement. Even so, their clarification requires some effort. This writer regards the picture of a mixture of definite orientations as unrealistic because the choice of alternate orientations is not unique. This is shown, for example, by considering the analogous problem of γ-ray polarization in the phenomenon of two-photon positron annihilation. No polarization of the γ rays is observed with a single detector A, sensitive to polarization but which receives only one photon from any pair. Now, operate A in coincidence with another detector B sensitive to linear polarization and actuated by the other photon of a pair; A will detect a partial linear polarization. If, on the other hand, B is sensitive to circular polarization, A detects a partial circular polarization. Having found that the polarization observed at A depends qualitatively on observations performed at B, one has little basis for classifying the "polarization of individual photons" detected by A, as either linear or circular.

The difficulties met in these discussions derive in part from the circumstance that quantum mechanics has been mostly concerned with the "pure" states of systems, such as states of definite spin orientation, which are represented by state vectors ψ. States with less than maximum information, represented by density matrices ρ, have been considered primarily in statistical mechanics and their discussion has been influenced by the historical background in this field. Moreover, most atomic physicists have not had much opportunity to become familiar with the representation of states by density matrices.[1]

Calculations can be, and usually have been, performed without reference to density matrices even when maximum information is not available. This is done by calculating as though a maximum of variables were observed and then summing (or averaging) over the eigenstates of the unobserved variables. On the other hand, when the states are represented by density matrices one can often expedite the calculations and avoid the introduction of unnecessary variables, which is of particular value for the treatment of many-body problems.

The variables to be dispensed with include the arbitrary phase of the state vector ψ which represents the state of the whole system.

It is also important that the density matrix of a system can be easily expressed in terms of the mean values of observables. Identifying a state by means of such physical parameters brings out the operational basis of the theory and helps in forming a mental picture.

In the course of a phenomenon, the variations of density matrices, or of equivalent sets of parameters, can be represented and analyzed to a considerable extent by operator techniques without reference to special representations of the operators. The whole treatment of quantum-mechanical problems in terms of density matrices can thereby reflect the features of physical phenomena more directly and in closer correspondence to macroscopic methods than is otherwise possible.

This review proposes to familiarize the reader with the methods and the potentialities of quantum-mechanical treatments in terms of density matrices, utilizing a variety of examples. It starts by stating briefly the basic facts regarding the identification of states for atomic systems and regarding the properties of density matrices as developed by von Neumann (N27).* The treatment is then developed toward the systematic application of operator techniques. Most, if not all, of the factual material is well known, but the point of view departs in some respects from that of standard references (T38, N55).

The analysis of an experimental situation in Sec. 5 is designed to illustrate why partially polarized light should be described not as a mixture of different polarizations, but by parameters that define its observable properties without reference to fictitious models.

This review does not cover the field of quantum statistical mechanics, but it may clarify initial concepts and approaches to this field.

2. IDENTIFICATION OF STATES

(a) Pure States

Quantum mechanics deals usually with phenomena in which a maximum of information is available about the system under consideration. This maximal information is attained, for example, for a spinless particle whose state at a certain time is represented by a wave function $\psi(x,y,z)$, for a nonrelativistic spinning electron in a central field with a specified full set of quantum numbers (n,l,j,m), or for a light beam of intensity I, frequency ν, direction $\boldsymbol{\omega}$, and linear polarization $\boldsymbol{A}$. States of maximal information are often called "pure states" or simply "states."

A pure state is characterized by the existence of an experiment that gives a result predictable with certainty when performed on a system in that state and in that state only. For example, linear polarization of a light beam in a given plane is characterized by 100% transmission of *each photon* through a suitably oriented Nicol prism; no other state of polarization is fully transmitted by the same prism. Filtration through a Nicol prism defines a state of polarization completely because beams thus filtered behave identically with respect to any other polarization analyzer.

An experiment that yields a unique predetermined result for a system in a given pure state can be designed to act as a filter which leaves the system undisturbed, like a Nicol prism traversed by light of the pertinent linear polarization. The experiment may then be repeated again and again on the same system, at least in principle, always with certainty as to its outcome. Pure states can, in fact, be "prepared" by subjecting systems to a filter-type experiment. The example of light polarization often proves particularly helpful to understand the relationships between quantum mechanical states and the experiments that characterize them; the relationships may be less obvious in other examples but are essentially the same.

An experiment that characterizes uniquely a pure state, as indicated above, and thus provides a maximum of information about it is called a "complete" experiment. The term "complete" may have a relative meaning, with reference to only a part of the variables of a system. For example, filtration through a Nicol prism is a complete experiment with regard to polarization, but has little relevance to the photon energy.

A pure state can then be identified by specifying the complete experiment that characterizes it uniquely. Mathematically one can construct a variety of Hermitian operators which have the given pure state as an eigenstate. An Hermitian operator which transforms the states of a physical system represents an observable.[2]

---

[1] The name "statistical matrix" is often used instead of "density matrix." The name density matrix itself relates to the correspondence between ρ and the distribution function $\rho(q_i,p_i)$ in the phase space of classical statistical mechanics. This correspondence has been developed by Wigner (W32), see Sec. 9.

* References are listed at the end of this article.

[2] This statement is believed valid for most applications in quantum mechanics but is subject to restrictions discussed in (WWW52).

Given such an operator, it proves possible to design, at least in principle, an experiment that constitutes a measurement of the corresponding observable. For example the state (n,l,j,m) of an electron bound in a central field is an eigenstate of the energy operator H corresponding to the eigenvalue E_{nlj} and also an eigenstate of the angular momentum operator J_z corresponding to the eigenvalue $m\hbar$. Experiments that measure energies and angular momenta are familiar; experiments to measure less familiar observables can nevertheless be designed.

When it is not convenient to identify a pure state by specifying the relevant complete experiment or its corresponding operator, the state may be identified as a linear superposition of eigenstates of any suitable complete set of operators. For example the wave function $\psi(x,y,z)$ of a particle describes its state as a superposition of all simultaneous eigenstates of the three position operators x, y and z, the amplitude of the eigenstate (x,y,z) being $\psi(x,y,z)$. The representation of a pure state, either as an eigenstate of a particular operator or as a superposition of eigenstates of another arbitrary operator, is usually called a "state vector" ψ. (It is also often called a wave function even when ψ is not expressed in terms of space coordinates.)

(b) General

Quantum-mechanical systems also occur for which no complete experiment gives a unique result predictable with certainty. For example, no polarization analyzer admits or rejects with certainty photons of partially polarized light. We say, loosely, that the information on such a system is less than a maximum, with reference to the lack of a complete experiment with a uniquely predetermined outcome. The state of the system is nevertheless fully identified by any data adequate to predict the (statistical) results of all conceivable observations on the system. Whether or not the predicted dispersion of these statistical results attains its theoretical minimum is irrelevant to the concept of state. Indeed "state" means whatever information is required about a specific system, in addition to physical laws, in order to predict its behavior in future experiments.

States that are not "pure" have been called "mixed" states because they can be described by the incoherent superposition of pure states. Incoherent superposition means, by definition, that to calculate the probability of finding a certain experimental result with a system in the mixed state one must first calculate the probability for each of the pure states and then take an average, attributing to each of the pure states an assigned "weight."

The concept of a nonpure state as a mixture of pure states originates from the early investigations of the connections between quantum mechanics and statistical mechanics. Classical statistical mechanics deals, ideally, with statistical mixtures (*ensembles*) of classic mechanical states, each of which is characterized by sharply defined values of all variables. The pure states of quantum mechanics were then visualized as analogues of the classic mechanical states and quantum-statistical mechanics was regarded as dealing with mixture of pure states.

However, the description of a nonpure state as the incoherent superposition of pure states is not unique. There is in general no reason, for example, why unpolarized light should be described as a mixture of two linear polarizations rather than of two circular ones.

We shall regard the state of a system, whether pure or not, as defined by its previous history, i.e., by the method of its preparation. Information on the preparation may be replaced by adequate experimentation on an ensemble of identically prepared systems. Fragmentary information, inadequate to make statistical predictions about future experiments, is often complemented by plausible assumptions. Examples given in Sec. 4 show how the information is represented by a density matrix. It will also be shown that lack of information about a variable can often be expressed as a statement that certain operators have expectation value zero. For example if the spin orientation of a particle is wholly unknown, the expectation value of each component of its angular momentum vanishes.

The point of view taken here differs to some extent from that which is often followed in standard treatments of statistical quantum mechanics, but the difference may be regarded as a matter of preference rather than of substance. To illustrate this difference, it may be recalled how Tolman (T38) analyzes the fluctuations in the experimental results obtained from a system in a given state. He considers first a subset of fluctuations that should be expected if the system were in any one of a number of pure states. Then he combines these subsets into a broader set according to the probability distribution that the system be in any of these pure states. In this paper we consider only the broader set of all fluctuations among the experimental results obtained with an ensemble of systems prepared according to identical specifications. We do not analyze these fluctuations into subsets when this analysis is not unique and does not correspond to observable characteristics of the situation. We regard as somewhat incidental whether a change in the preparation of the systems would reduce the fluctuations to a lower level or whether they already attain the minimum level set by quantum-mechanical laws. That is, we deal with a *single statistical ensemble* of quantum mechanical systems *prepared by identical procedures*, not with a statistical ensemble of quantum mechanical ensembles.[3] The terms "mean" or "average"

---

[3] The point of view of Elsasser (E37) is in some respects close to that adopted in this paper, but differs rather substantially in other respects, because it rests on the notion that "there exists no general *a priori* principle of how to collect samples which can be represented by the same statistical matrix" and that accordingly "the problem of determining the statistical matrix is essentially an indefinite one." Here we take a more positive attitude toward the determination of the statistical matrix.

will be always understood, in this review, to relate to the statistical expectation regarding such an ensemble.

3. DENSITY MATRIX

(a) Formulas for Mean Values

In the ordinary formalism of quantum mechanics, a pure state is identified by the coefficients c_n of the expansion of its state vector ψ into eigenvectors u_n of some complete set of operators,

$$\psi=\sum_n c_n u_n. \tag{3.1}$$

For a system in this state, an operator Q represented by a matrix $Q_{n'n}$ has the mean value

$$\langle Q\rangle=\sum_{n'n} Q_{n'n} c_{n'}^{*} c_n. \tag{3.2}$$

When a nonpure state is represented by the incoherent superposition of a number of pure states $\psi^{(i)}$ with statistical weights $p^{(i)}$, to each pure state corresponds a mean value $\langle Q\rangle_i$, and the mean value of Q for the incoherent superposition is given by the grand average

$$\langle Q\rangle=\sum_i p^{(i)}\langle Q\rangle_i=\sum_{nn'} Q_{n'n}\sum_i p^{(i)} c^{(i)}{}_{n'}{}^{*} c^{(i)}{}_n. \tag{3.3}$$

One defines then the density matrix as

$$\rho_{nn'}=\sum_i p^{(i)} c^{(i)}{}_{n'}{}^{*} c^{(i)}{}_n \tag{3.4}$$

so that (3.3) becomes

$$\langle Q\rangle=\sum_{n'n} Q_{n'n}\rho_{nn'}=\sum_{n'}(Q\rho)_{n'n'}=\mathrm{Tr}(Q\rho), \tag{3.5}$$

where $\mathrm{Tr}A$ indicates the trace (i.e., the sum of the diagonal elements) of a matrix A.

(b) Definition

We regard the density matrix as defined by Eq. (3.5) rather than by (3.4). It represents a minimum set of input data which serves to calculate the mean value of any operator Q for a system prepared according to given specifications.[4] The information from which this set of data is derived is equivalent to a knowledge of the mean values of as many independent operators $Q^{(r)}$ as there are independent parameters in the matrix $\rho_{nn'}$. In fact the initial information on the state of the system is often conveniently expressed as a set of $\langle Q^{(r)}\rangle$ from which the $\rho_{nn'}$ are determined by solving a system of Eq. (3.5), one for each $\langle Q^{(r)}\rangle$. Equation (3.5) which gives $\langle Q\rangle$ for a generic Q in terms of the matrix ρ may be looked upon as a device to calculate a generic $\langle Q\rangle$ from advance knowledge of a special set of $\langle Q^{(r)}\rangle$'s.

(c) Limitations

Limitations on the matrix elements $\rho_{nn'}$ include the following:

(1) The condition that $\langle Q\rangle$ is real for every Hermitian operator Q, requires ρ to be *Hermitian* too,

$$\rho_{n'n}=\rho_{nn'}^{*}. \tag{3.6}$$

(2) The condition that the unit operator I has the mean value 1, requires

$$\mathrm{Tr}(I\rho)=\mathrm{Tr}(\rho)=\sum_n\rho_{nn}=1. \tag{3.7}$$

(3) The condition that every operator with non-negative eigenvalues (for example the operator $\delta_{kn}\delta_{kn'}$ with eigenvalues 1 and 0) has a non-negative mean value, requires ρ to be *positive definite.* That is, every diagonal element of ρ in any matrix representation must be non-negative,

$$\rho_{kk}\geqslant 0, \tag{3.8}$$

The mean value of the operator $\delta_{kn}\delta_{kn'}$, namely ρ_{kk}, represents the probability of finding the system by a suitable experiment in the pure state u_k. Accordingly, $\mathrm{Tr}\rho=\sum_k\rho_{kk}$ represents the total probability of finding the system in any one of a complete set of orthogonal states and the condition (3.7) constitutes an obvious requirement.

(4) The Hermitian matrix ρ may be reduced to the diagonal form

$$\rho_j\delta_{jj'}=\sum_{nn'} T_{jn}\rho_{nn'}T^{-1}{}_{n'j'} \tag{3.9}$$

by a unitary transformation T. The conditions (3.7) and (3.8) require that

$$\sum_j\rho_j^2\leq\left(\sum_j\rho_j\right)^2=[\mathrm{Tr}(\rho)]^2=1. \tag{3.10}$$

The $\sum_j\rho_j^2$ is the $\mathrm{Tr}(\rho^2)$, and therefore we have in general

$$\mathrm{Tr}(\rho^2)=\sum_{nn'}|\rho_{nn'}|^2\leq 1, \tag{3.11}$$

which limits the value of every single element of the density matrix.

(d) The Variations in Time

The variations in time of the density matrix are governed by the Schroedinger equation. Any unitary transformation S that changes the vector ψ representing a pure state into $S\psi$, also changes the matrix ρ representing a generic state into $S\rho S^{-1}$. The ordinary Schroedinger equation $\partial\psi/\partial t=-i\hbar^{-1}H\psi$ has, as a formal solution, the time variable transformation $\psi(t)=\exp(-i\hbar^{-1}Ht)\psi(0)$. The corresponding formulas for the density matrix are

$$\partial\rho/\partial t=-i\hbar^{-1}[H\rho-\rho H], \tag{3.12}$$

$$\rho(t)=\exp(-i\hbar^{-1}Ht)\rho(0)\exp(i\hbar^{-1}Ht). \tag{3.13}$$

The mean values of observables vary in time according to

$$\begin{aligned}\langle Q\rangle_t=\mathrm{Tr}[Q\rho(t)]&=\mathrm{Tr}[Q\exp(-i\hbar^{-1}Ht)\rho(0)\exp(i\hbar^{-1}Ht)]\\&=\mathrm{Tr}[\exp(i\hbar^{-1}Ht)Q\exp(-i\hbar^{-1}Ht)\rho(0)]\\&=\mathrm{Tr}[Q(t)\rho(0)].\end{aligned} \tag{3.14}$$

---

[4] Any prediction about the behavior of a system can be expressed as the mean value of a suitable operator Q. For example, the probability of a certain event is the mean value of an operator whose eigenvalues are 1 when the event occurs and 0 when it does not.

The transition from the Schroedinger representation ($Q=\text{const}$, $\rho=\rho(t)$) to the Heisenberg representation ($Q=Q(t)$, $\rho=\text{const}$) appears here as a trivial consequence of the identity $\text{Tr}(AB)=\text{Tr}(BA)$.

In the scheme where the Hamiltonian H is diagonal, with eigenstates v_m, (3.13) takes the form

$$\rho_{mm'}(t)=\rho_{mm'}(0)\exp[i\hbar^{-1}(E_{m'}-E_m)t]. \tag{3.15}$$

Notice that only energy differences and the corresponding frequencies $\hbar^{-1}(E_{m'}-E_m)$ appear here, to the exclusion of absolute energies which are indeed not observable.

(e) The Diagonal Representation

The diagonal representation (3.9) of the density matrix is equivalent to

$$\rho_{nn'}=\sum_j \rho_j T_{jn'}{}^* T_{jn}. \tag{3.16}$$

Comparison of (3.16) with (3.4) shows that a generic state can be represented by the incoherent superposition of orthogonal pure states ψ_j with statistical weights ρ_j, the state vectors ψ_j being identified by their components which are the elements of the transformation matrix T_{jn}. This representation is *not unique* in general. That is, an infinity of different representations (3.4) can be constructed for a given density matrix unless: (a) one stipulates that the states $\psi^{(i)}$ represented by the $c_n{}^{(i)}$'s are mutually orthogonal, like the ψ_j's, and (b) no two nonzero eigenvalues ρ_j are equal. Examples of the relevance of these conditions will be given in the following sections.

(f) A Pure State

A pure state is represented by a density matrix with one eigenvalue equal to 1 and all others equal to 0. (The eigenvalues of the matrix are invariant under unitary transformations). Equation (3.16) reduces then to $\rho_{nn'}=T_{1n'}{}^*T_{1n}$. A density matrix $\rho_{nn'}$ can be factorized in this manner when, and only when, all minors of its determinant vanish.

(g) The Number of Independent Parameters

The number of independent parameters that identify a density matrix depends on its number N of rows and columns. This is the number of orthogonal pure states over which the $\sum_n$ extends in (3.1) and in the subsequent formulas. This number may be infinite but is often finite when one considers only a particular property of a system, such as the spin orientation of a nucleus.

A matrix with N rows and columns has N^2 elements. The condition (3.6), that ρ be Hermitian, restricts the number of independent real parameters in the N^2 complex elements to N^2, and (3.7) restricts it further to N^2-1. These parameters are further limited in their range of variation by inequalities like (3.8) and (3.11), but are nevertheless independent. Therefore *it takes*, in general, *N^2-1 separate measurements*, or equivalent data, *to identify the state* of a system which possesses N independent pure states. For example, a nucleus of Na^{23} with spin $\frac{3}{2}$ has 4 independent pure states of spin orientation and the description of a specific state of orientation requires in general 15 independent data.[5]

Note, for comparison, that a pure state is identified by the N complex coefficients c_n of (3.1). The number of meaningful independent real parameters in the c_n's is reduced from $2N$ to $2N-1$ by the normalization condition $\sum_n c_n{}^2=1$, and further to $2N-2$ because the phase of the state vector ψ is arbitrary and physically meaningless.

(h) Interacting Systems

The joint state of two interacting systems a and b is represented by a density matrix $\rho^{(ab)}$ of which each row or column is labelled by two indices (m,n) corresponding to eigenstates u_m and v_n of the separate systems.[6] An operator $Q^{(a)}$ of the system a may be treated as an operator of the whole system when multiplied by the unit operator of b, $\mathbb{1}^{(b)}=\delta_{n'n}$. The mean value of $Q^{(a)}$ is then

$$\begin{aligned}\langle Q^{(a)}\rangle&=\text{Tr}(Q^{(a)}\mathbb{1}^{(b)}\rho^{(ab)})\\&=\sum_{mm'} Q_{m'm}{}^{(a)}\Big[\sum_{nn'}\delta_{n'n}\rho_{mn,m'n'}{}^{(ab)}\Big]\\&=\text{Tr}_a(Q^{(a)}\rho^{(a)}),\end{aligned} \tag{3.17}$$

where

$$\rho_{mm'}{}^{(a)}=\sum_n \rho_{mn,m'n}{}^{(ab)}=[\text{Tr}_b(\rho^{(ab)})]_{mm'} \tag{3.18}$$

represents the information on a alone contained in $\rho^{(ab)}$.

The states of the two systems are *uncorrelated* when

$$\langle Q^{(a)}Q^{(b)}\rangle=\langle Q^{(a)}\rangle\langle Q^{(b)}\rangle \tag{3.19}$$

for all pairs $Q^{(a)}Q^{(b)}$ which requires the joint density matrix to be a product,

$$\rho_{mn,m'n'}{}^{(ab)}=\rho_{mm'}{}^{(a)}\rho_{nn'}{}^{(b)}. \tag{3.20}$$

4. EXAMPLES

(a) Orientation of Spin-1/2 Particles

This spin orientation is represented by a density matrix with two rows and columns, corresponding to

---

[5] The clause "in general" refers to the possibility that a few initial measurements turn out to be compatible only with certain special values of the remaining parameters. In this event further measurements of these parameters would be unnecessary. For example, if a first experiment on an assembly of Na^{23} nuclei utilizes a Stern-Gerlach inhomogeneous magnet and if every nucleus is found to follow the path corresponding to the magnetic quantum number $m=\frac{1}{2}$, the state is already fully identified.

[6] The two systems may consist of two portions of matter, of two different groups of particles (like nuclei and electrons of the same portion of matter) or of different characteristics (like spin and orbital motion) of the same particle. One can regard a system as an aggregate of two separate systems whenever there exists a complete set of operators O, pertaining to the whole system, that is the sum of two sets of operators $O^{(a)}$ and $O^{(b)}$, such that $O^{(b)}$ does not operate on the eigenstates u_m of $O^{(a)}$, and $O^{(a)}$ does not operate on the v_n's of $O^{(b)}$.

two pure states of opposite spin orientation (e.g., up and down). In a generic state the degree and direction of spin orientation are indicated by the magnitude and direction of the vector

$$\mathbf{P}=\langle\boldsymbol{\sigma}\rangle=\mathrm{Tr}(\rho\boldsymbol{\sigma}) \tag{4.1}$$

whose components are the mean values of the operators represented by the three Pauli matrices σ_x, σ_y, and σ_z. A pure state with definite spin orientation has $P=1$, a state of random orientation has $P=0$. For particles with magnetic moment μ in a magnetic field $\mathbf{H}$, under conditions of paramagnetic polarization, we have $\mathbf{P}=\mu\mathbf{H}/\frac{1}{2}kT$.

Knowledge of $\mathbf{P}$ is sufficient to identify the density matrix, considering: (1) that any 2×2 Hermitian matrix can be represented as a linear combination of σ_x, σ_y, σ_z and of the unit matrix $\mathbb{1}$, (2) that the Pauli matrices have the properties $\mathrm{Tr}\sigma_i=0$, $\mathrm{Tr}(\sigma_i\sigma_k)=2\delta_{ik}$, (3) that $\mathrm{Tr}\rho=1$ and $\mathrm{Tr}\mathbb{1}=2$ require the coefficient of $\mathbb{1}$ to be $\frac{1}{2}$. We have, then, (TG49)

$$\rho=\tfrac{1}{2}(\mathbb{1}+P_x\sigma_x+P_y\sigma_y+P_z\sigma_z)=\tfrac{1}{2}(\mathbb{1}+\mathbf{P}\cdot\boldsymbol{\sigma})$$
$$=\tfrac{1}{2}\begin{vmatrix} 1+P_z & P_x-iP_y \\ P_x+iP_y & 1-P_z \end{vmatrix}. \tag{4.2}$$

The representation (4.2) leads to a simple treatment of the Larmor precession of spin orientation in a magnetic field. The magnetic moment $\boldsymbol{\mu}$ is represented quantum-mechanically by the operator $\gamma(\frac{1}{2})\hbar\boldsymbol{\sigma}$, where γ is the gyromagnetic ratio, so that the Hamiltonian in a magnetic field $\mathbf{H}$ is $-\gamma(\frac{1}{2})\hbar\boldsymbol{\sigma}\cdot\mathbf{H}$. The Schroedinger equation (3.12) is then

$$\partial\rho/\partial t=\tfrac{1}{2}\partial\mathbf{P}/\partial t\cdot\boldsymbol{\sigma}=i\tfrac{1}{4}\gamma[\mathbf{H}\cdot\boldsymbol{\sigma}\;\mathbf{P}\cdot\boldsymbol{\sigma}-\mathbf{P}\cdot\boldsymbol{\sigma}\;\mathbf{H}\cdot\boldsymbol{\sigma}]$$
$$=-\tfrac{1}{2}\gamma\mathbf{H}\times\mathbf{P}\cdot\boldsymbol{\sigma}, \tag{4.3}$$

and thus reduces to the classical equation

$$\partial\mathbf{P}/\partial t=-\gamma\mathbf{H}\times\mathbf{P}. \tag{4.4}$$

The extension of this equation to spin values $j>\frac{1}{2}$ is indicated below, the extension to include the paramagnetic relaxation is given in Sec. 11g.

The response of a particle detector, whose efficiency depends on spin orientation, and which therefore serves as a polarization analyzer, may be represented by an operator in a form analogous to (4.2). Maximum and minimum efficiency, ϵ_M and ϵ_m, correspond necessarily, for spin $\frac{1}{2}$, to particles with opposite spin orientations indicated by unit vectors $\mathbf{Q}$ and $-\mathbf{Q}$. The detector is then represented by an operator D with the spin orientation eigenstates $\mathbf{Q}$ and $-\mathbf{Q}$ and the eigenvalues ϵ_M and ϵ_m. When the z axis is parallel to $\mathbf{Q}$, the matrix of this operator is

$$D=\begin{vmatrix} \epsilon_M & 0 \\ 0 & \epsilon_m \end{vmatrix}, \tag{4.5}$$

and for a generic orientation of axes it is

$$D=\tfrac{1}{2}[(\epsilon_M+\epsilon_m)\mathbb{1}+(\epsilon_M-\epsilon_m)\mathbf{Q}\cdot\boldsymbol{\sigma}]. \tag{4.6}$$

The probability of response of this detector to particles with the density matrix (4.2) is the mean value

$$\langle D\rangle=\mathrm{Tr}(\rho D)=\tfrac{1}{2}[(\epsilon_M+\epsilon_m)+(\epsilon_M-\epsilon_m)\mathbf{P}\cdot\mathbf{Q}]. \tag{4.7}$$

(b) Orientation of Spin-j Particles

The density matrix of spin orientation has $2j+1$ rows and columns. For $j>\frac{1}{2}$ the representation (4.2) can be generalized to include, besides the unit matrix and terms depending on dipole polarization, additional terms which depend on multipole polarization (see end of Sec. 6). This expansion of ρ constitutes a "reduction" in the sense of tensor algebra (FR57, Sec. 19) and the coefficients of the operators may be called the "reduced elements" of ρ.[7] The application of the expansion to collision and disintegration processes will be outlined in Sec. 11 a.

The definition (4.1) of the spin orientation vector is replaced for a generic j by

$$\mathbf{P}=\langle\mathbf{J}\rangle/j\hbar, \tag{4.8}$$

where $\mathbf{J}$ indicates the angular momentum operator. With this definition, the changes of spin orientation induced by a magnetic field $\mathbf{H}$ are described by (4.4) for all values of j.[8,9] The multipole parameters of spin orientation remain constant, under the influence of $\mathbf{H}$, provided they are expressed in a coordinate system which precesses with $\mathbf{P}$.

(c) Polarization of Electromagnetic Radiation

Light polarization is represented by a density matrix with two rows and columns, corresponding to two opposite polarizations, e.g. to linear polarizations indicated by orthogonal unit vectors $\mathbf{A}_1$ and $\mathbf{A}_2$. Because of mathematical analogy to the density matrix for the orientation of spin-$\frac{1}{2}$ particles, the density matrix of light polarization can be represented in the form (4.2), but the indices x, y, and z no longer correspond to directions of the physical space. They relate to a mathematical representation of polarizations in a 3-dimen-

---

[7] These coefficients have also been called "statistical tensors" (F51a) and "state multipoles" (F53).

[8] This result is derived from the Schroedinger Eq. (3.12), taking into account that: (1) the magnetic moment $\boldsymbol{\mu}$ is represented by the operator $\gamma\mathbf{J}$, (2) therefore, the Hamiltonian is $-\gamma\mathbf{J}\cdot\mathbf{H}$, (3) the definition (4.8) requires the dipole term in the expansion of ρ to have the form $3\mathbf{P}\cdot\mathbf{J}[j(j+1)(2j+1)\hbar]^{-1}$, (4) the commutator $[\mathbf{J}\cdot\mathbf{H},\mathbf{P}\cdot\mathbf{J}]$ equals $i\hbar\mathbf{H}\times\mathbf{P}\cdot\mathbf{J}$.

[9] The history of (4.4) has some interest. The treatment of spin precession by ordinary quantum-mechanical methods is very complicated (BR45). Bloch pointed out, in his paper on nuclear induction (B46), that (4.4) can be established without solving the Schroedinger equation, because the quantum-mechanical mean value of any quantity follows in its time dependence exactly the classical equations of motion. Later Wangsness and Bloch (WB53) derived (4.4) from the Schroedinger equation by a density matrix procedure equivalent to that indicated here. The link between the classical and the Schroedinger equation emerges clearly when ρ is expressed in terms of mean values, as in (4.2), so that the Schroedinger equation for ρ reduces to an equation of motion for mean values. A general formulation of this reduction will be given in Sec. 6.

sional space, called "Poincaré representation," in which the z axis corresponds to linear polarization along $\mathbf{A}_1$, negative z to polarization along $\mathbf{A}_2$, positive x to linear polarization at 45° between $\mathbf{A}_1$ and $\mathbf{A}_2$ and positive y to circular polarization rotating from $\mathbf{A}_1$ toward $\mathbf{A}_2$. A unitary transformation of the base polarizations $\mathbf{A}_1$ and $\mathbf{A}_2$ is represented by a rotation of Cartesian axes in the representative space.

This representation of the density matrix has proved convenient (F49, F54a), in the same manner as the corresponding representation for spinning particles. The form (4.2) of the density matrix is normalized to provide the input for the calculation of elementary processes involving one photon at a time. For macroscopic processes, one multiplies (4.2) simply by the intensity I of a radiation beam. The quantities I, IP_x, IP_y, IP_z constitute a set of Stokes parameters which characterizes simultaneously the intensity and polarization state of the beam (S52, F49, FMcD51).

(d) Joint Polarization of Photon Pairs

In the positron annihilation by two-photon process, which was mentioned in Sec. 1, the emitted γ rays display a polarization only when the two photons of a pair are analyzed simultaneously but not when the photons are observed singly. The experimental results can be described in terms of the mean value $\langle D^{(A)}D^{(B)}\rangle$, where the operators D represent polarization analyzers serving to detect two photons A and B in coincidence. The matrix of $D^{(A)}$ or $D^{(B)}$ is given by (4.6), where $\mathbf{Q}$ indicates now a vector of the Poincaré representation of polarizations. The usual statement that "the photons of a pair have opposite polarizations" means actually that, for perfect analyzers with $\epsilon_m=0$, one would find $\langle D^{(A)}D^{(B)}\rangle=(1/4)\epsilon_M{}^{(A)}\epsilon_M{}^{(B)}(1-\mathbf{Q}^{(A)}\cdot\mathbf{Q}^{(B)})$, which vanishes for $\mathbf{Q}^{(A)}=\mathbf{Q}^{(B)}$ and equals $(\frac{1}{2})\epsilon_M{}^{(A)}\epsilon_M{}^{(B)}$ for opposite analyzer settings $\mathbf{Q}^{(A)}=-\mathbf{Q}^{(B)}$, whether $\mathbf{Q}^{(A)}$ corresponds to linear, circular, or elliptical polarization. In a realistic application with imperfect analyzers one finds

$$\langle D^{(A)}D^{(B)}\rangle=\tfrac{1}{4}[(\epsilon_M+\epsilon_m)^{(A)}(\epsilon_M+\epsilon_m)^{(B)} - (\epsilon_M-\epsilon_m)^{(A)}(\epsilon_M-\epsilon_m)^{(B)}\mathbf{Q}^{(A)}\cdot\mathbf{Q}^{(B)}]. \tag{4.9}$$

The joint state of polarization of the photon pairs is represented by a density matrix $\rho^{(AB)}$ with 4 rows and columns, which can always be expressed as a linear combination of the 16 matrices $\mathbb{1}^{(A)}\mathbb{1}^{(B)}$, $\mathbb{1}^{(A)}\sigma_x{}^{(B)}$, $\cdots\sigma_z{}^{(A)}\sigma_z{}^{(B)}$. The requirement that $\mathrm{Tr}(\rho^{(AB)}D^{(A)}D^{(B)})$ be given by (4.9) for all choices of ϵ and $\mathbf{Q}$ implies that

$$\rho^{(AB)}=\tfrac{1}{4}(\mathbb{1}^{(A)}\mathbb{1}^{(B)}-\boldsymbol{\sigma}^{(A)}\cdot\boldsymbol{\sigma}^{(B)}). \tag{4.10}$$

This matrix represents the situation in very condensed form, without reference to any specific type of polarization. The negative sign corresponds to the statement that the "photons have *opposite* polarization." The density matrix for a single photon is, according to (3.18), $\rho^{(A)}=\mathrm{Tr}_B\rho^{(AB)}=(\frac{1}{2})\mathbb{1}^{(A)}$ and indicates natural polarization as expected.

The joint density matrix (4.10) can be derived, of course, by a theoretical calculation of the annihilation process.

(e) Steady States (Constant in Time)

When the state of a system does not vary in the course of time, the density matrix is constant. According to (3.15) we have then, in the scheme of energy eigenstates,

$$\rho_{mm'}=0 \quad \text{for} \quad E_m\neq E_{m'}. \tag{4.11}$$

Any operator Q of the system may be split into two parts, $Q=Q^{(c)}+Q^{(v)}$, one constant and one variable, by defining

$$Q_{mm'}{}^{(c)}=Q_{mm'}\delta_{E_mE_{m'}}, \quad Q_{mm'}{}^{(v)}=Q_{mm'}(1-\delta_{E_mE_{m'}}). \tag{4.12}$$

A steady state may then be characterized by the property

$$\langle Q^{(v)}\rangle=0 \quad \text{for all } Q. \tag{4.13}$$

In the Heisenberg representation formalism, $Q^{(v)}$ may be defined by $Q^{(v)}(t)=Q(t)-Q(0)$.

(f) States with Space Invariances

Space invariance may be treated like time invariance. For example if the state of a system is invariant under translations along the x axis, its density matrix has a property analogous to (4.11) in the scheme of momentum eigenstates, namely

$$\rho_{p_xp_{x'}}=0 \quad \text{for} \quad p_x\neq p_{x'}. \tag{4.14}$$

States invariant under space rotation, i.e., states of random orientation and spherical symmetry, are very important. Here the relevant scheme is that of eigenstates of the squared angular momentum $\mathbf{J}^2$ and of one of its components, e.g., J_z. Spherical symmetry is represented by

$$\rho_{\alpha jm,\,\alpha'j'm'}=0 \quad \text{for} \quad j\neq j' \quad \text{or} \quad m\neq m', \tag{4.15}$$

where j and m relate, as usual, to the eigenvalues of $\mathbf{J}^2$ and J_z; α is an additional quantum number relating to the eigenvalues of any additional operators that commute with $\mathbf{J}^2$, J_z and among themselves.

In the more restricted case of axial symmetry only, we have

$$\rho_{\alpha jm,\,\alpha'j'm'}=0 \quad \text{for} \quad m\neq m'. \tag{4.16}$$

A nucleus that had initially random spin orientation and has absorbed, or interacted with, a beam of unpolarized radiation has axial symmetry about the direction of the beam.[10] Any operator Q can be split, as in (4.12), into a

---

[10] This property is fundamental for the theory of angular correlations. It was first stated tentatively, in a different form, by Falkoff and Uhlenbeck (FU50) and then proved by various authors (L50, S50, L51, TG51), through arguments equivalent to the elementary considerations which lead to (4.16).

part $Q^{(s)}$ having a symmetry property of interest, e.g. invariance under space rotation, and a residual part $Q^{(\text{non-}s)}$ which is not invariant. The mean value $\langle Q^{(\text{non-}s)}\rangle$ vanishes for symmetric states.

(g) States of Thermal Equilibrium

Statistical mechanics shows that the state of a system at a temperature T is represented by the incoherent superposition of eigenstates of energy E_m with weights proportional to the Boltzmann distribution factor $\exp(-E_m/kT)$. In order that the sum of the weights for all eigenstates equal 1, the weight of each state must equal the factor $\exp(-E_m/kT)$ divided by the "sum of states" $Z(T)=\sum_m \exp(-E_m/kT)$. That is, the density matrix is diagonal in the scheme of energy eigenstates and is given by

$$\rho_{mm'}=\frac{e^{-E_m/kT}}{Z(T)}\delta_{mm'}, \tag{4.17}$$

in this scheme and by

$$\rho=\frac{e^{-H/kT}}{Z(T)}=\frac{e^{-H/kT}}{\mathrm{Tr}(e^{-H/kT})} \tag{4.18}$$

in operator notation.

Because ρ is a function of the Hamiltonian operator only, and this function can be expanded in series, the mean value of any operator Q can be expressed in terms of the traces $\mathrm{Tr}(QH^n)$. In particular we have

$$\langle Q^{(i)}\rangle=0 \tag{4.19}$$

for the whole class of operators $Q^{(i)}$ such that $\mathrm{Tr}(Q^{(i)}H^n)=0$ for all n, i.e. such that $\mathrm{Tr}[Q^{(i)}f(H)]=0$ for all functions of H. The derivation of (4.18) can be inverted to show that, whenever (4.19) holds, ρ is a function of H only.

The functional dependence of ρ on H can be identified by the mean values $\langle H\rangle$, $\langle H^2\rangle\cdots\langle H^n\rangle\cdots$, which are given by well known formulas of statistical mechanics in terms of the derivatives of $Z(T)$, e.g., $\langle H\rangle=-d\ln Z/d(1/kT)$, $\langle H^2\rangle=\langle H\rangle^2+[d/d(1/kT)]^2\ln Z$, etc.

The derivation of (4.18) through statistical mechanics is, then, equivalent to showing that, for a system in contact with a thermal reservoir, the mean values of all operators $Q^{(i)}$ tend to zero in the course of time and the $\langle H^n\rangle$ approach the values predicted by thermodynamics. These results should follow from the Wangsness-Bloch equation discussed in Sec. 11g.

5. ANALYSIS OF LIGHT EMITTED BY ATOMS

As a further illustration we discuss here the polarization of light emitted by atoms under controlled conditions in an ideal experimental arrangement (F49) (Fig. 1). The emission takes place in a chamber C on which an optical detector and polarization analyzer are collimated from a variable direction. The atoms enter the chamber as a molecular beam, preanalyzed by a Stern-Gerlach device A to select atoms with a specified magnetic quantum number m, referred to an axis z parallel to the magnetic field in A. To be specific, we consider an optical transition from a p state ($l=1$) to a d state ($l'=2$) and assume that the Stern-Gerlach analyzer A has removed all atoms with $m\neq 0$. In the ordinary language of quantum mechanics we say, then, that the transition will lead from ($l=1$, $m=0$) to ($l'=2$, $m'=0, \pm 1$).

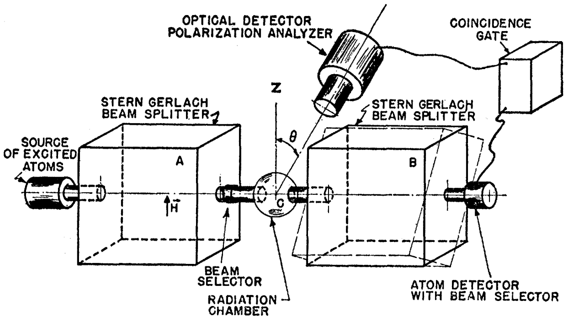

FIG. 1. Block diagram of an ideal experiment.

The angular distribution of intensity and polarization of the light emitted by these atoms can be calculated, for example, by considering the light as the incoherent superposition of three components emitted respectively in transitions to $m'=1$, 0 and -1. Alternately, one can derive by a density matrix treatment a general formula (FR57, 19.6) for the intensity and polarization distribution of radiation emitted in transitions to a state whose orientation is not observed.

The result can be expressed as a density matrix for the polarization, in the form (4.2), multiplied by an intensity distribution. Both the density matrix and the intensity distribution are functions only of the angle θ between the direction of emission and the z axis, since the state of atoms filtered by A has cylindrical symmetry about this axis if the atom velocity is disregarded. The result is

$$I(\theta)\rho\propto(1+\tfrac{1}{6}\sin^2\theta)\left[1+\frac{\sin^2\theta}{6+\sin^2\theta}\sigma_\zeta\right], \tag{5.1}$$

where the subscript ζ indicates the direction in the representative Poincaré space corresponding to linear polarization in a plane containing the z axis. In explicit matrix notation with the first row and column referring to this polarization and the second to linear polarization orthogonal to z, we have

$$I(\theta)\rho\propto\left|\begin{matrix}1+(4/3)\sin^2\theta & 0\\ 0 & 1\end{matrix}\right|. \tag{5.2}$$

Because the matrix is diagonal in this scheme, the light can be regarded as the incoherent superposition of two beams, one polarized linearly in a vertical plane, with intensity $\propto 1+(4/3)\sin^2\theta$ and one polarized in a horizontal direction with intensity $\propto 1$. The excess polarization in the vertical plane relates to the fact that the transition to $m'=0$, associated with an oscillating current parallel to z, is more likely than the transition to either $m'=\pm 1$ by a factor 4/3. However this analysis of the polarization provides little justification for regarding the light as "a mixture of photons" with the two linear polarizations. That the density matrix is diagonal in the scheme (5.2) follows from the symmetry of the initial state $m=0$ under rotations about z and under reflection on the xy plane.

If the device A selects atoms in the initial state $m=1$, instead of $m=0$, the symmetry under reflection on the xy plane drops out and the density matrix of the polarization is more complicated. We have here

$$I(\theta)\rho \propto \left(1-\frac{1}{14}\sin^2\theta\right)\left[\mathbb{1}+\frac{10\cos\theta\,\sigma_\eta-\sin^2\theta\,\sigma_\zeta}{14-\sin^2\theta}\right]$$

$$=\left(1-\frac{1}{14}\sin^2\theta\right)\left[\mathbb{1}+\frac{(100\cos^2\theta+\sin^4\theta)^{\frac{1}{2}}}{14-\sin^2\theta}\times(\cos\alpha\,\sigma_\eta-\sin\alpha\,\sigma_\zeta)\right], \quad (5.3)$$

where the subscript η refers to circular polarization and the angle $\alpha=\mathrm{arctg}(\sin^2\theta/10\cos\theta)$ indicates the type of partial polarization, which varies from left circular to elliptical to linear and again to elliptical and right circular as θ varies from 0° to 180°.

Consider now another Stern-Gerlach device B added at the exit of the radiation chamber C, which can select atoms having a specified quantum number m' *after* the light emission process. These atoms are to be detected at the exit of B and the detector output can be used to "gate" the light analyzer device. The light will then appear to be completely polarized. If the magnetic fields in B and A are parallel and A selects $m=0$, atoms will pass through B in coincidence with radiation emission for $m'=1$, 0 or -1. For $m'=0$ the polarization is linear, in a vertical plane; for $m'=\pm 1$ it is circular, elliptic or linear depending on the direction of observation. The intensity and polarization distribution for each m' is given by

$$I(\theta)\rho \propto \tfrac{2}{3}\sin^2\theta(\mathbb{1}+\sigma_\zeta), \quad \text{for } m'=0$$

$$I(\theta)\rho \propto \tfrac{1}{4}(1+\cos^2\theta)\left(\mathbb{1}-\frac{\sin^2\theta\sigma_\zeta\mp 2\cos\theta\sigma_\eta}{1+\cos^2\theta}\right), \quad \text{for } m'=\pm 1. \quad (5.4)$$

The sum of the three intensity and polarization distributions yields (5.1).

The elementary treatment leads one to visualize the partially polarized light represented by (5.1) as a mixture of three kinds of photons polarized, respectively, according to (5.4), simply because it appears natural to classify the final state according to the same axis of quantization as the initial state. (Notice that we have here *three* incoherent components, whereas the diagonal matrix (5.2) suggests only two.) However, this representation does not rest on any intrinsic property of the emitted light. Indeed, *if the analyzer B is tilted* about the axis of the atomic beam, so that the final quantum number m' relates to an axis oblique to z, the beam of atoms after radiation will be resolved into *five* components, with $m'=0$, ± 1, ± 2. Each of these components is associated with a light emission fully polarized, so that the light is in fact resolved into *five* incoherent polarized components, when the analyzer is gated by the output of B. Each of the five components is represented by a matrix $I(\theta)\rho$ and the sum of these matrices equals (5.1).

In summary, light is completely polarized only when maximum information regarding the state of orientation of the atoms which have emitted it (or have interacted with it) is provided, and actually taken into account. The information is required regarding the orientation both before and after the process. Incomplete information leads to partial polarization. Different representations of partial polarization as the incoherent superposition of polarized components correspond to different conceivable methods of completing the information on the atoms. None of these alternate representations appears intrinsically relevant—apart from analytical convenience—unless the complete information on the atom is actually collected and correlated with the polarization analysis.

An experimental analysis, equivalent to that indicated here for light polarization, can be developed for the orientation of particle spins or for any other quantum-mechanical variable.

We have dealt here specifically with the information or lack of information on the orientation of the atoms *after* light emission, that is, with a follow-up analysis. One could as well have considered alternate types of pre-analysis. For example one could arrange the selective device A so that it first splits up the atom beam components with different m and then refocuses them together on the entrance of the radiation chamber. It is then possible to select one or the other m value at will by stopping the other beam components with a diaphragm, or else to allow two or more beam components to radiate together.

6. EXPANSION IN ORTHOGONAL OPERATORS

The density matrix of the orientation of spin-$(\frac{1}{2})$ particles is decomposed by (4.2) into a sum of standard matrices $\mathbb{1}$ and $\boldsymbol{\sigma}$. This decomposition was found convenient. It replaces, in effect, the elements of the density matrix with the coefficients of the standard matrices, namely 1, P_x, P_y, and P_z in (4.2), which are directly observable parameters being the mean values of physical quantities. Indeed (4.4) constitutes an equation of motion of the spin expressed entirely in terms of the observable vector parameter **P**.

This representation can be generalized. The operators $\mathbb{1}$, σ_x, σ_y, and σ_z are orthogonal, in the sense that the product of any two of these operators has trace zero. For a generic system one may consider a set of operators U_i which are Hermitian, except when otherwise noted, and obey the orthogonality condition

$$\mathrm{Tr}(U_iU_k)=C\delta_{ik}. \tag{6.1}$$

C represents a normalization factor which will be taken as 1, for convenience, unless otherwise noted. For systems whose complete operators have N orthogonal eigenstates, so that ρ has N rows and columns, N^2 linearly independent operators constitute a complete orthogonal set of base operators.

Each operator Q of the system can be expanded into a sum of U_i's,

$$Q=\sum_i \mathrm{Tr}(QU_i)U_i=\sum_i Q_iU_i. \tag{6.2}$$

In particular the density matrix is represented by

$$\rho=\sum_i \mathrm{Tr}(\rho U_i)U_i=\sum_i\langle U_i\rangle U_i=\sum_i \rho_iU_i. \tag{6.3}$$

The state is thereby identified by the parameters ρ_i which are the mean values $\langle U_i\rangle$ of the base operators. The mean value of any operator Q is a linear combination of the ρ_i's,

$$\langle Q\rangle=\mathrm{Tr}(Q\rho)=\sum_i Q_i\rho_i. \tag{6.4}$$

Suitable choice of the U_i's serves to eliminate irrelevant variables, much like the choice of coordinates does in classical physics. The initial state of a system is often identified by the fact that certain types of operators have mean value zero, as seen in the examples of Sec. 4. It is then convenient to include these operators among the U_i's. With regard to the final state of the system one is often interested only in predicting the results of certain experiments, i.e. the mean values of only a few operators. If these operators are included in the U_i's, then only a few of the ρ_i's have to be calculated. This approach is most effective, of course, when the equations of motion interlink only some subsets of the ρ_i's.

It is usually convenient that one of the base operators be the unit operator, to within a normalization constant. This operator will be indicated as

$$U_0=[\mathrm{Tr}(\mathbb{1})]^{-\frac{1}{2}}\mathbb{1}, \tag{6.5}$$

where the normalization is such that $C=1$ in (6.1). The operator U_0 commutes with all other operators, and its mean value

$$\rho_0=\langle U_0\rangle=[\mathrm{Tr}(\mathbb{1})]^{-\frac{1}{2}} \tag{6.6}$$

is the same for all states of the system, owing to the normalization condition $\mathrm{Tr}\{\mathbb{1}\rho\}=1$. A state is then identified by the N^2-1 parameters ρ_i with $i\neq 0$. The condition (6.1) requires that

$$\mathrm{Tr}\{U_0U_i\}=\mathrm{Tr}\{\mathbb{1}U_i\}=\mathrm{Tr}\{U_i\}=0,\quad \text{for}\quad i\neq 0, \tag{6.7}$$

i.e., that the sum of the eigenvalues of U_i vanish for $i\neq 0$. States of random spin orientation or of zero polarization, i.e., states of minimum information, are represented by a density matrix proportional to $\mathbb{1}$ ("equiprobability") and, therefore, by zero values of ρ_i for $i\neq 0$. Departures of ρ_i from zero represent positive elements of information about the state.

When there are N rows and columns, $U_0=N^{-\frac{1}{2}}\mathbb{1}$. When N is infinitely large, the situation is somewhat akin to that encountered in the normalization of wave functions in the continuum and in the application of the Dirac δ function.

The Schroedinger equation (3.12) takes the form of a system of linear equations among the mean values $\rho_i=\langle U_i\rangle$. Replacing ρ with $\sum_k \rho_kU_k$ on the right of (3.12), multiplying (3.12) by U_i and taking the trace one obtains

$$\begin{aligned}\partial\rho_i/\partial t&=-i\hbar^{-1}\sum_k \mathrm{Tr}\{U_i[H,U_k]\}\rho_k\\ &=i\hbar^{-1}\sum_k \mathrm{Tr}\{H[U_i,U_k]\}\rho_k,\end{aligned} \tag{6.8}$$

that is

$$\partial\rho_i/\partial t=\sum_k \Omega_{ik}\rho_k, \tag{6.9}$$

where Ω_{ik} is an antisymmetric matrix. The set of commutators $[U_i,U_k]$ can again be expanded into a sum of base operators

$$i[U_i,U_k]=\sum_p c_{ik}{}^pU_p,\quad c_{ik}{}^p=\mathrm{Tr}\{U_pi[U_i,U_k]\}, \tag{6.10}$$

where the coefficients $c_{ik}{}^p$, antisymmetric in (i,k) and real, are an important property of the set of base operators.

The matrix Ω_{ik} of the system of Eqs. (6.9) is therefore real,

$$\Omega_{ik}=\sum_p c_{ik}{}^p\,\mathrm{Tr}\{\hbar^{-1}HU_p\}=-\Omega_{ki}, \tag{6.11}$$

and represents an infinitesimal orthogonal transformation. Notice that $\Omega_{i0}=\Omega_{0k}=0$ since all U_i's commute with U_0. The eigenvalues of Ω are imaginary and equal to i times the beat frequencies $\hbar^{-1}(E_{m'}-E_m)$ among the eigenvalues E_m of H, as seen in (3.15).

The operators U_i need not be Hermitian. If they are not, one must consider also their adjoint, or Hermitian conjugate, operators $U_i^+=\tilde{U}_i^*$ (where $\sim$ indicates transposition and * complex conjugation). Equations (6.1) to (6.4) are then replaced by

$$\mathrm{Tr}\{U_iU_k^+\}=\delta_{ik}, \tag{6.12}$$

$$Q=\sum_i \mathrm{Tr}\{QU_i^+\}U_i=\sum_i Q_i^+U_i=\sum_i Q_iU_i^+, \tag{6.13}$$

$$\rho=\sum_i \mathrm{Tr}\{\rho U_i^+\}U_i=\sum_i \rho_i^+U_i=\sum_i \rho_iU_i^+, \tag{6.14}$$

$$\langle Q\rangle=\sum_i Q_i^+\rho_i=\sum_i Q_i\rho_i^+, \tag{6.15}$$

and (6.10) by

$$\begin{aligned}&i[U_i,U_k^+]=\sum_p c_{ik}{}^pU_p,\\ &\qquad c_{ik}{}^p=\mathrm{Tr}\{U_p^+i[U_i,U_k^+]\}=-c_{ki}{}^{p*},\end{aligned} \tag{6.16}$$

whereas (6.9) and (6.11) remain unchanged.

The immediate generalizations of the set of Pauli matrices, $\mathbb{1}$, σ_x, σ_y, and σ_z include: (a) the set of 16 double operators for two-photon polarization $\mathcal{I}^{(A)}\mathcal{I}^{(B)}$, $\mathcal{I}^{(A)}\sigma_x^{(B)}, \cdots, \sigma_z^{(A)}\sigma_z^{(B)}$, utilized in Sec. 4d, (b) the sets of 16 Dirac operators $\mathcal{I}$, γ_1, $\gamma_2 \cdots$, $\gamma_1\gamma_2\gamma_3\gamma_4$ or $\mathcal{I}$, ρ_1, $\rho_1\sigma_x \cdots \rho_3\sigma_z$, (c) the sets of $(2j+1)^2$ multipole operators pertaining to the orientation of particles with spin j, which were referred to in Sec. 4b. The multipole operators are conveniently chosen in tensorial form U_{kq}, i.e. so that they transform under coordinate rotations like the spherical harmonic functions $Y_{kq}(\theta,\varphi)$, with k integer running up to $2j$ and q running in integral steps from $-k$ to k (F53). The $2j+1$ operators with a given k and different q represent the components of the 2^k-pole moment of the particle, to within a normalization constant. Because of the Wigner-Eckart theorem (FR57, Sec. 14), the matrices of the multipole operators are standard Wigner coefficients, properly normalized.

The operators $Q^{(i)}$, defined in Sec. 4g as orthogonal to all powers of the Hamiltonian H, may also conveniently be chosen as orthogonal to one another; together with a set of orthogonal functions of H, they form a complete system. The Wigner phase-space density operators to be discussed in Sec. 9 also constitute a complete orthogonal set.

Orthogonal sets of operators have a role analogous in many respects to the role of orthogonal coordinates in the phase space of classical particle mechanics.[11]

7. HILBERT SPACE REPRESENTATION

Equations (6.4), and (6.2) and (6.3) as well, have the familiar form of scalar products of vectors. A geometrical representation in Hilbert space is indeed appropriate but this space is not the usual Hilbert space of quantum mechanics.

In the usual representation the N eigenstates u_n of (3.1) are taken as base vectors, and all pure states ψ are vectors of unit length of an N-dimensional Hilbert space. The density matrix $\rho_{nn'}$ is a second degree tensor of this space.

Here we consider a Hilbert space with N^2 dimensions, of which the eigenstate products $u_n u_{n'}^+$ constitute one set of base vectors. The density matrix ρ is a *vector* in this space, with the components $\rho_{nn'}$ in the coordinate system of base vectors $u_n u_{n'}^+$. Any operator Q is also represented by a vector, with the components $Q_{n'n}$ in this system, and the mean value $\langle Q\rangle$ may be visualized, according to (3.5), as a scalar product $\mathbf{Q}\cdot\boldsymbol{\rho}$.

The condition (6.1), with $C=1$, defines the set of operators U_i as an orthonormal system of base vectors, which replaces the system $u_n u_{n'}^+$. The following equations of Sec. 6 simply restate the vector equations of the Hilbert space in terms of the new coordinate systems. In vector notation (6.2), (6.3), and (6.4) read, respectively, $\mathbf{Q}=\sum_i(\mathbf{Q}\cdot\mathbf{U}_i)\mathbf{U}_i=\sum_i Q_i\mathbf{U}_i$, $\boldsymbol{\rho}=\sum_i(\boldsymbol{\rho}\cdot\mathbf{U}_i)\mathbf{U}_i=\sum_i \rho_i\mathbf{U}_i$, and $\langle Q\rangle=\boldsymbol{\rho}\cdot\mathbf{Q}$. All physical equations are real when expressed in terms of the hermitian operators U_i.

Equation (6.6), equivalent to (3.7), fixes the value of the component ρ_0 of $\boldsymbol{\rho}$, and thus states that the tip of the vector which represents $\boldsymbol{\rho}$ as drawn from the origin of the Hilbert space lies on the hyperplane of points $\mathbf{x}$ with the coordinate $x_0=[\mathrm{Tr}(\mathcal{I})]^{-\frac{1}{2}}=N^{-\frac{1}{2}}$. The condition (3.11) takes the form $\boldsymbol{\rho}\cdot\boldsymbol{\rho}\leq 1$ and thus limits the tip of $\boldsymbol{\rho}$ to the interior of a hypersphere of radius 1. The condition $\boldsymbol{\rho}\cdot\boldsymbol{\rho}=1$ is achieved only for pure states, of which there is only a $(2N-2)$-dimensional manifold, as discussed in Sec. 3g. In mathematical language, this manifold at which $\boldsymbol{\rho}\cdot\boldsymbol{\rho}$ reaches 1 consists of the "extreme points" of the "convex core" which contains all possible positions of the tip of $\boldsymbol{\rho}$ (W32, p. 79).

The unitary transformations S of the pure state vectors ψ are represented by all rotations of the ordinary Hilbert space of quantum mechanics. In the Hilbert space considered here the S are also rotations, which transform a vector $\mathbf{Q}$ representing the matrix Q into a vector $\mathbf{Q}'$ representing SQS^{-1}. The S are only a *subgroup* of all rotations in this space, because they are limited by the requirement of leaving invariant the manifold of extreme points $\boldsymbol{\rho}\cdot\boldsymbol{\rho}=1$. The ordinary Hilbert space is convenient for ordinary quantum mechanical applications because it represents as generic rotations the transformations S, which include the changes of state arising in the course of time from conservative interactions within a system. The representation considered here is convenient for more general applications, in particular for the transformations arising from interactions that are in effect nonconservative (see Secs. 10 and 11).

8. MEASURES OF INFORMATION

Pure states of a system, with $\boldsymbol{\rho}\cdot\boldsymbol{\rho}=\mathrm{Tr}(\rho^2)=1$ afford a maximum of information; states with $\rho=\rho_0 U_0=[\mathrm{Tr}(\mathcal{I})]^{-1}\mathcal{I}$, and $\boldsymbol{\rho}\cdot\boldsymbol{\rho}=[\mathrm{Tr}(\mathcal{I})]^{-1}=N^{-1}$, afford a minimum of information. The statistical fluctuations of a physical quantity, represented by $\langle Q^2\rangle-\langle Q\rangle^2=\mathbf{Q}\cdot\boldsymbol{\rho}-(\mathbf{Q}\cdot\boldsymbol{\rho})^2$, tend to increase as $\boldsymbol{\rho}\cdot\boldsymbol{\rho}$ decreases from 1 to N^{-1}, because the second term is quadratic in $\boldsymbol{\rho}$.

The eigenvalues ρ_j have the distribution $(1,0,0\cdots)$ for pure states and become increasingly uniform as $\boldsymbol{\rho}\cdot\boldsymbol{\rho}$ decreases, to the point of being all equal to N^{-1} for states of minimum information. This transition of the spectrum of ρ may be represented by variations of functions of the form

$$\textstyle\sum_j f(\rho_j)=\mathrm{Tr}[f(\rho)] \tag{8.1}$$

which serve as indices of the quantity of information about the system. A quantity of this type is independent of the representation of ρ and also invariant in the course of time if ρ obeys the Schroedinger equation.

The quantity

$$\textstyle\sum_j \rho_j^2=\mathrm{Tr}(\rho^2)=\boldsymbol{\rho}\cdot\boldsymbol{\rho}\leq 1 \tag{8.2}$$

---

[11] The description of states and of their variations in terms of standard operators and of their mean values can be developed in many directions. For example, Bopp (B56) has been considering standard operators with positive eigenvalues, with the aim of treating the variations of their positive mean values as stochastic processes.

serves itself well as a measure of information, for many purposes. In the representation in terms of orthogonal operators (Sec. 6), we have

$$\mathrm{Tr}(\rho^2)=\sum_{i=0}^{N^2-1}\rho_i^2=N^{-1}+\sum_{i=1}^{N^2-1}\rho_i^2=N^{-1}+\sum_{i=1}^{N^2-1}\langle U_i\rangle^2. \quad (8.3)$$

The quantity of information above the lower limit N^{-1} appears here to consist of additive contributions represented by the squared mean values of operators with trace zero. The spin operators σ of (4.2), the operators $Q^{(v)}$ of (4.12) and $Q^{(i)}$ of (4.18) are all traceless. The degree of polarization P in the spin orientation density matrix (4.2), or in the corresponding light polarization matrix, is closely related to $\mathrm{Tr}(\rho^2)$, since in this example we have

$$\mathrm{Tr}(\rho^2)=\tfrac{1}{2}+\tfrac{1}{2}P^2,\quad P^2=2\sum_i\left\langle\frac{\sigma_i}{2}\right\rangle^2. \quad (8.4)$$

One may generalize the concept of degree of polarization by defining

$$P^2=\frac{N}{N-1}\sum_{i=1}\rho_i^2. \quad (8.5)$$

Qualitatively, the idea of quantity of information implies that the total information in a number of separate systems is the sum of the quantities of information on the various systems. According to (3.20), the density matrix $\rho^{(ab)}$ of two uncorrelated systems is the product of density matrices $\rho^{(a)}$ and $\rho^{(b)}$ of the two systems. We have then

$$\mathrm{Tr}(\rho^{(ab)2})=\mathrm{Tr}(\rho^{(a)})^2\,\mathrm{Tr}(\rho^{(b)})^2. \quad (8.6)$$

Thus $\ln \mathrm{Tr}(\rho^2)$ is additive and corresponds to the concept of quantity of information better than $\mathrm{Tr}(\rho^2)$ itself. The $\ln \mathrm{Tr}(\rho^2)$ ranges from $-\ln N$ up to 0.

In statistical mechanics it has been known for a long time (N55, T38), that the entropy of a system coincides with the mean value of $-k\ln\rho$,[12] where k is the Boltzmann constant, that is, with

$$\langle -k\ln\rho\rangle=-k\,\mathrm{Tr}(\rho\ln\rho). \quad (8.7)$$

The connection between entropy and quantity of information suggests then that $\mathrm{Tr}(\rho\ln\rho)=\langle\ln\rho\rangle$, or, still better, $\mathrm{Tr}[\rho\ln(N\rho)]=\langle\ln(N\rho)\rangle$, be regarded as a suitable definition for the quantity of information. This definition allows one to introduce a *quantity of information operator* $\ln(N\rho)$, which has itself the additivity property

$$\ln(N_{ab}\rho^{(ab)})=\ln(N_a\rho^{(a)})+\ln(N_b\rho^{(b)}) \quad (8.8)$$

whenever the states of a and b are uncorrelated.

Notice that $\mathrm{Tr}(\rho\ln\rho)$ and $\ln\mathrm{Tr}(\rho^2)$ vary between the same limits and that their values are never very different. Since $\mathrm{Tr}(\rho^2)$ may be expressed as $\langle\rho\rangle$, we are comparing here $\langle\ln\rho\rangle$ with $\ln\langle\rho\rangle$.

The condition $\mathrm{Tr}(\rho\ln\rho)=\text{minimum}$ may be utilized in quantum mechanics much like the condition of maximum entropy in thermodynamics or the condition of maximum number of complexions in classical statistical mechanics. The condition $\mathrm{Tr}(\rho\ln\rho)=\min$ with the subsidiary condition that certain operators $L^{(r)}$ have assigned mean values $\langle L^{(r)}\rangle$ determines (N27, S55) that

$$\rho=\exp(\gamma-\textstyle\sum_r\gamma_r L^{(r)}). \quad (8.9)$$

The constants γ, γ_r are fixed as usual by the subsidiary conditions. The state of thermal equilibrium (4.17) is obtained by this procedure, when the set of $L^{(r)}$ consists of the Hamiltonian H only; then $\gamma_1=1/kT$ and $\gamma=-\ln Z$. The discussion of Sec. 4g shows that the condition $\mathrm{Tr}(\rho\ln\rho)=\min$ incorporates into ρ the following information, in addition to the values of the $\langle L^{(r)}\rangle$: (a) all quantities $Q^{(i)}$ such that $\mathrm{Tr}[Q^{(i)}f(L^{(1)}\cdots L^{(r)}\cdots)]=0$ have mean value zero, (b) all quantities $f(L^{(1)}\cdots L^{(r)}\cdots)$ have mean values which are specified functions of the $\langle L^{(r)}\rangle$ and of the parameters γ, γ_r.

9. WIGNER PHASE SPACE DENSITY OPERATORS

Wigner has defined (W32) a quantum analog of the function $\rho(\mathbf{R},\mathbf{P})$ which represents a statistical distribution in phase space of the states of a particle according to classical mechanics. If the state of the particle is identified by the density matrix in the scheme of position coordinates $\mathbf{r}$, $(\mathbf{r}'|\rho|\mathbf{r}'')$, the Wigner distribution function is

$$\rho(\mathbf{R},\mathbf{P})=\frac{1}{(\hbar\pi)^3}\int d\mathbf{r}(\mathbf{R}+\mathbf{r}|\rho|\mathbf{R}-\mathbf{r})\exp(i\hbar^{-1}\mathbf{P}\cdot 2\mathbf{r}). \quad (9\ 1)$$

This function can be generalized immediately to a many-particle system and has proved useful in statistical mechanics (see, e.g., M48, IZ51).

The values of the function $\rho(\mathbf{R},\mathbf{P})$ constitute a set of parameters $\rho_i=\langle U_i\rangle$, as defined by (6.3). They are the mean values of the operators $U_{\mathbf{RP}}$ with the matrix representation

$$(\mathbf{r}'|U_{\mathbf{RP}}|\mathbf{r}'')=(\hbar\pi)^{-3}\delta(2\mathbf{R}-\mathbf{r}'-\mathbf{r}'')\times\exp[i\hbar^{-1}\mathbf{P}\cdot(\mathbf{r}'-\mathbf{r}'')]. \quad (9.2)$$

($\mathbf{R}$ and $\mathbf{P}$ are parameters, not operators.) The operators $U_{\mathbf{RP}}$ are Hermitian, have trace h^{-3}, and form an orthogonal set, since one verifies from (9.2) that

$$\mathrm{Tr}(U_{\mathbf{R}'\mathbf{P}'}U_{\mathbf{RP}})=h^{-3}\delta(\mathbf{R}-\mathbf{R}')\delta(\mathbf{P}-\mathbf{P}'), \quad (9.3)$$

but this set does not include the unit operator. The normalization constant C of (6.1) is, therefore, h^{-3}. Any operator with a matrix $(\mathbf{r}'|Q|\mathbf{r}'')$ can be represented as a linear combination of $U_{\mathbf{RP}}$; the coefficients $h^3\,\mathrm{Tr}(QU_{\mathbf{RP}})$ are calculated by selecting out the matrix elements of Q on the skew diagonal $\mathbf{r}'+\mathbf{r}''=2\mathbf{R}$ and taking their Fourier transform.

---

[12] The matrix $\ln\rho$ is identified by having the same eigenstates as ρ and the eigenvalues $\ln\rho_j$, equal to the logarithm of the eigenvalues of ρ. Stratonovich (S55) has emphasized that $\mathrm{Tr}(\rho\ln\rho)$ is conveniently expressed as $\sum_j\rho_j\ln\rho_j$ even when ρ is initially identified otherwise than in terms of its eigenvalues and eigenfunctions.

The operator $U_{\mathbf{RP}}$ applied to a function $f(\mathbf{r})$ changes it into $(\hbar\pi)^{-3}\exp[i\hbar^{-1}\mathbf{P}\cdot 2(\mathbf{r}-\mathbf{R})]f(2\mathbf{R}-\mathbf{r})$. It follows that $U_{\mathbf{RP}}$ has eigenfunctions of the form

$$f(\mathbf{r})=\exp(i\hbar^{-1}\mathbf{P}\cdot\mathbf{r})u^{(\pm)}(\mathbf{r}-\mathbf{R}), \tag{9.4}$$

where $u^{(\pm)}$ is an even or odd function of its variable. The corresponding eigenvalues are $\pm(\hbar\pi)^{-3}$, respectively. These eigenvalues are highly degenerate; complete systems of orthogonal functions (9.4) can be constructed.[13]

The eigenfunctions (9.4) are wave packets which represent states of the particle with $\langle\mathbf{r}\rangle=\mathbf{R}$ and $\langle\mathbf{p}\rangle=\mathbf{P}$. The existence of negative eigenvalues of $U_{\mathbf{RP}}$ constitutes a quantum effect to be contrasted with the condition $\rho(\mathbf{R},\mathbf{P})\geq 0$ obeyed by the classical distribution function.

10. INTERACTION WITH AN UNOBSERVED SYSTEM. IRREVERSIBILITY

When two systems a and b interact, or have been in interaction for a certain period of time, practical interest often centers in the resulting state of a only, irrespective of what has become of b. This state is described at a time t, according to (3.18) and (3.13), by

$$\begin{aligned}(\rho^{(a)})_t&=\mathrm{Tr}_b[(\rho^{(ab)})_t]\\&=\mathrm{Tr}_b[\exp(-i\hbar^{-1}Ht)(\rho^{(ab)})_{t=0}\exp(i\hbar^{-1}Ht)],\end{aligned} \tag{10.1}$$

where the Hamiltonian may be regarded as

$$H=H_a+H_b+V, \tag{10.2}$$

V being the interaction.

The density matrix $\rho^{(a)}$, as defined by (10.1), does not obey the Schroedinger Eq. (3.12) or the equivalent Eq. (6.9). If orthonormal operator sets U_i and V_r are introduced for the systems a and b, respectively, the complete density matrix $\rho^{(ab)}$ is identified by parameters ρ_{ir}, according to (6.3), $\rho^{(ab)}=\sum_{ir}\rho_{ir}U_iV_r$, and $\rho^{(a)}$ by the subset of parameters ρ_{i0},

$$\rho^{(a)}=[\mathrm{Tr}_b(\mathbb{1})]^{\frac{1}{2}}\sum_i\rho_{i0}U_i. \tag{10.3}$$

The Schroedinger Eq. (6.9) governs the simultaneous variation of all the ρ_{ir}, but we are interested here in the variation of the ρ_{i0} only. In the Hilbert space representation, the complete density matrix is represented by a vector $\boldsymbol{\varrho}^{(ab)}$ which experiences a rotation in the course of time. The components ρ_{i0} define, through (10.3), a vector $\boldsymbol{\varrho}^{(a)}$ which is the projection of $\boldsymbol{\varrho}^{(ab)}$ on the subspace with base vectors U_iV_0. The vector magnitude $\boldsymbol{\varrho}^{(ab)}\cdot\boldsymbol{\varrho}^{(ab)}$ is conserved in the course of time but the magnitude of its projection, $\boldsymbol{\varrho}^{(a)}\cdot\boldsymbol{\varrho}^{(a)}$, which indicates the quantity of information on the system a, is *not conserved*, in general.

In problems of atomic collision often one does not attempt to follow the time variations of $(\rho^{(a)})_t$ but one inquires about the density matrix of the final state f, corresponding to $t=\infty$. The operator $\exp[-i\hbar^{-1}Ht]$ in (10.1) takes then an asymptotic value for $t=\infty$, which may be represented by a scattering or reaction matrix T and (10.1) becomes

$$(\rho^{(a)})_f=\mathrm{Tr}_b(\rho^{(ab)})_f=\mathrm{Tr}_b[T(\rho^{(ab)})_0T^{-1}]. \tag{10.4}$$

If one calculates $(\rho^{(a)})_t$ or $(\rho^{(a)})_f$ as a matrix, one takes only the trace over eigenstates of b, as indicated in (10.1) and (10.4). Alternately one may want to calculate the corresponding parameters $(\rho_{i0})_t$ or $(\rho_{i0})_f$, or even only the mean value of certain operators $Q^{(a)}$ which are of interest. In this event one takes the trace over eigenstates of both b and a,

$$\begin{aligned}(\rho_{i0})_t=[\mathrm{Tr}_b(\mathbb{1})]^{-\frac{1}{2}}\,\mathrm{Tr}_{ab}[U_i\exp(-i\hbar^{-1}Ht)(\rho_{ab})_0\\\times\exp(i\hbar Ht)],\end{aligned} \tag{10.5}$$

$$\begin{aligned}\langle Q^{(a)}\rangle=\mathrm{Tr}_{ab}[Q^{(a)}\exp(-i\hbar^{-1}Ht)(\rho_{ab})_0\\\times\exp(i\hbar Ht)].\end{aligned} \tag{10.6}$$

Equation (10.1) is a suitable point of departure for the description of irreversible processes, such as viscosity or the relaxation of spin orientation in nuclear induction. Quantum mechanics is set up to treat reversible, conservative, processes, and deals in principle with any phenomenon by extending the treatment to a system large enough to be isolated, and hence conservative. Indeed the complete system $a+b$ considered in this section is assumed to be isolated and the variations of $(\rho^{(ab)})_t$ in the course of time are reversible, inasmuch as the initial state description $(\rho^{(ab)})_0$ can be generated mathematically from $(\rho^{(ab)})_t$ by the inverse transformation $\exp(i\hbar^{-1}Ht)(\rho^{(ab)})_t\exp(-i\hbar^{-1}Ht)$.

In the calculation of $(\rho^{(a)})_t$ according to (10.1) much information about the whole system is discarded when one calculates the $\mathrm{Tr}_b[\cdots]$. It is nevertheless implied, in general, that the full matrix $(\rho^{(ab)})_t$, which results from a reversible transformation, has to be calculated prior to the trace calculation. Furthermore, the calculation of $(\rho^{(ab)})_t$ implies, at least in principle, a calculation of $(\rho^{(ab)})_{t'}$ for each t' between 0 and t.

Irreversible processes are special phenomena in which much of the information included in $\rho^{(ab)}$ at a time t' is not relevant to $\rho^{(a)}$ not only at the same instant t' but also at later times $t>t'$. For example, when a nuclear spin $\mathbf{S}_i$, part of a macroscopic amount of matter, is driven to precess by a magnetic field, its precession exerts a magnetic disturbance on the surrounding particles and these disturb, in turn, particles further away. Owing to the very large number of particles, the initial disturbance keeps propagating away and will not react back onto $\mathbf{S}_i$ at later times to any significant extent. The details of this disturbance may accordingly be ignored. That is, one may replace at each time t' the actual $(\rho^{(ab)})_{t'}$ with a simpler $(\rho_s{}^{(ab)})_{t'}$ which contains less information about b without introducing any significant error in the final result $(\rho^{(a)})_t$ of (10.1). In the Hilbert space representation, the actual vector $\boldsymbol{\varrho}^{(ab)}$ is

---

[13] Dr. M. S. Green points out that the operator $U_{\mathbf{RP}}$ is represented, in terms of projection operators $\mathcal{P}_{\mathbf{R}}{}^{(\pm)}$ such that $\mathcal{P}_{\mathbf{R}}{}^{(\pm)}f(r)=(1/2)[f(\mathbf{r})\pm f(2\mathbf{R}-\mathbf{r})]$, by

$$U_{\mathbf{RP}}=(\hbar\pi)^{-3}\exp(i\hbar^{-1}\mathbf{P}\cdot\mathbf{r})[\mathcal{P}_{\mathbf{R}}{}^{(+)}-\mathcal{P}_{\mathbf{R}}{}^{(-)}]\exp(-i\hbar^{-1}\mathbf{P}\cdot\mathbf{r}).$$

replaced with a different, "shorter," $\boldsymbol{\varrho}_s^{(ab)}$ which not only has the same projection $\boldsymbol{\varrho}^{(a)}$ but also will evolve in the course of time so that its projection remains effectively the same as that of the correct $\boldsymbol{\varrho}^{(ab)}$. The difference between $\boldsymbol{\varrho}_s^{(ab)}$ and $\boldsymbol{\varrho}^{(ab)}$ is, however, essential from the standpoint of generating back $(\rho^{(ab)})_0$ by inverse transformation; $\exp(i\hbar^{-1}Ht)(\rho_s^{(ab)})_t\exp(-i\hbar^{-1}Ht)$ cannot be expected to resemble $(\rho^{(ab)})_0$. That is to say: the transformation of $(\rho_s^{(ab)})_t$ in the course of time is adequate to represent correctly $(\rho^{(a)})_t=\mathrm{Tr}_b(\rho^{(ab)})_t \sim\mathrm{Tr}_b(\rho_s^{(ab)})_t$ but it is definitely irreversible, whereas the transformation of $(\rho^{(ab)})_t$ is reversible.

In conclusion, the interaction of a with b constitutes an irreversible process when the resulting variation of $\rho^{(a)}$ may be adequately described in terms of an irreversible[14] $(\rho_s^{(ab)})_t$ instead of the correct $(\rho^{(ab)})_t$. Section 11g will indicate a simplification of (10.1) which constitutes an adequate approximation under typical conditions of irreversibility and which leads to a variation of $(\rho^{(a)})_t$ governed by a differential equation. Section 11h will give an equivalent but somewhat more systematic derivation of the same result.

11. APPLICATIONS OF OPERATOR TECHNIQUES

The results of quantum-mechanical calculations can always be expressed by the trace of multiple operator products, like (10.6). Whenever the initial or final states are not specified with a maximum of detail, the final formula of a calculation has in essence the same structure as (10.6), whichever analytical procedure has been employed.

The elementary procedure for calculating a $\mathrm{Tr}(AB\cdots F)$ is to calculate a matrix representation for each factor, A, B, $\cdots F$, and then to work out the multiple sum over products of matrix elements $\sum_{ij\cdots l} A_{ij}B_{jk}\cdots F_{hi}$. This procedure requires one to specify several states with complete sets of quantum numbers, indicated here by i, $j\cdots$, which are usually irrelevant to the problems. Therefore, this procedure not only is laborious but tends to obscure the underlying physical processes and the nature of approximation methods through the introduction of irrelevant variables.

These disadvantages can often be avoided by working on the formal solution of a problem, as indicated in Sec. 10, by a variety of operator techniques. An approach of this kind has been utilized widely in problems involving Dirac spinor variables (see, e.g., H54, p. 216) and is sometimes called the "Casimir trick" (C33, K55). Procedures of operator algebra and calculus (F51) and reduced matrix techniques (FR57) may be utilized in a variety of problems to take advantage of symmetries and invariances and to carry out approximation expansion, with no resort to the explicit matrix representation of operators. The correspondence to the classical variables represented by the operators remains thus in sight.

Many of these procedures have been developed recently and in connection with specialized applications. This section will survey a number of different examples in condensed form, referring to the original literature for details. The last two subsections, relating to irreversible processes with which the author has been particularly concerned, will be developed in somewhat greater detail.

(a) Rotational Invariance. Tensor Operators

For systems effectively isolated in space, the results of a calculation must be independent of the orientation of any coordinates which are utilized. As an example, we consider here the emission of a particle a by a radioactive nucleus which is left in a state b subject to no orientation analysis. We inquire about the angular and spin distribution of the radiation a, on the basis of information on the initial state of the nucleus indicated by *ab*.

A harmonic analysis of this distribution is performed, in effect, by representing $(\rho^{(a)})_t$ in the form $[\mathrm{Tr}_b\mathbb{1}]^{\frac{1}{2}}\sum_i \rho_{i0}U_i^+$, corresponding to (10.3) and (6.14), with the U_i^+ chosen to be *tensor operators*, i.e., such as to transform under coordinate rotations like the spherical harmonic functions $Y_{lm}(\theta,\varphi)$. (Because the Y_{lm} are complex, the transformation law implies that the U_i are not Hermitian, hence the expansion of $(\rho^{(a)})_t$ follows (6.14) instead of (6.3)). The parameters ρ_i vary contragrediently to the U_i^+ and are to be calculated through (10.5). In this equation the Hamiltonian H is invariant under coordinate rotations, if the nucleus is isolated. Since H is invariant and the whole $\mathrm{Tr}_{ab}[\cdots]$ in (10.5) is to be invariant, the initial state matrix $(\rho_{ab})_0$ may then contribute to the trace only through a portion of itself that varies under coordinate rotations contragrediently to U_i, so as to cancel out the variation of U_i. If $(\rho^{(ab)})_0$ is represented, according to (6.14), in the form $\sum_s \rho_s U_s^+$, with the U_s^+ chosen again to vary like spherical harmonics, there will usually be only one operator U_s^+ with $s=s_i$, that transforms contragrediently to U_i, and only this term $s=s_i$ will contribute to $\mathrm{Tr}_{ab}[\cdots]$. (At most there will be a few such terms.) Equation (10.5) reduces then to

$$(\rho_{i0})_t=[\mathrm{Tr}_b(\mathbb{1})]^{-\frac{1}{2}}\,\mathrm{Tr}_{ab}[U_i\exp(-i\hbar^{-1}Ht)U_{s_i} \times\exp(i\hbar^{-1}Ht)]\rho_{s_i}. \qquad (11.1)$$

Thus each parameter ρ_{i0} of the radiation distribution depends only on the parameter ρ_{s_i} of the initial nuclear state which has the same transformation law under coordinate rotations. This selection rule is qualitatively almost obvious. For example, if the initial nuclear state has axial symmetry—which happens when the nuclear

---

[14] Van Hove (VH55) has given a quantum mechanical treatment of irreversibility which has elements in common with the developments of Sec. 11g and h but which deals with a somewhat different question. It considers a single large system, such as a and b combined, and inquires about the variations of an operator Q which commutes with an unperturbed Hamiltonian, corresponding to H_a+H_b in (10.2), but not with V. The connections between the Van Hove approach and the approach presented in this review have not been fully explored.

spin is oriented by paramagnetic polarization—all its tensor parameters ρ_s that depend on coordinate rotations about the symmetry axis vanish. Equation (11.1) shows that the corresponding ρ_{i0}'s vanish too. Therefore the radiation distribution has also axial symmetry, which indeed was to be expected.

This analysis of calculations on the basis of rotational transformation properties has been exploited to various extents in applications to the theory of angular correlations and of disintegration processes (TG49, F51a, D52, WA52, CJ53, F53, W53, W54b), and is developed systematically by tensor algebra in (FR57, particularly Sec. 20).

(b) Depolarization Coefficients

An analysis based on tensor algebra also exploits the following consideration, for the calculation of $\mathrm{Tr}_{ab}[\cdots]$ in (11.1). Since the angular momentum is a constant of the motion, the Hamiltonian operator is diagonal in the scheme of eigenstates with quantum number j, where j is the spin of the initial nuclear state. It is then convenient to represent all operators in this scheme. The operators U_i of the emitted radiation are initially expressed in the scheme of angular momentum quantum numbers j_r of the radiation. The transformation from the scheme j_r to the scheme j depends on the spin j' of the residual nucleus, owing to the conservation of angular momentum in the disintegration, $\mathbf{J}=\mathbf{J}'+\mathbf{J}_r$, but involves only geometrical relations of coupling independent of the physical process of disintegration. When no observation is performed on the orientation of the residual spin, the transformation of U_i to the scheme j consists in essence of multiplying its matrix by a Racah coefficient $\bar{W}$ which is a function of the quantum numbers j, j_r, and j' and of the degree of the spherical harmonic with transforms like U_i. The factor $\bar{W}$ thus introduced in (11.1) is no larger than 1 in absolute value and thus reduces the value of the parameter $(\rho_{i0})_t$ which is a coefficient of the harmonic analysis of the radiation distribution. The larger the spin j' of the residual nucleus the smaller are $\bar{W}$ and $(\rho_{i0})_t$, so that the radiation distribution is more smeared out. This was also to be expected because more potential information has been lost by failing to observe the orientation of this nucleus. Therefore $\bar{W}$ may be called a *depolarization* coefficient. It indicates how much of the information originally available on the orientation of the nucleus fails to express itself in the radiation distribution owing to failure to observe the residual nucleus.

The experiment on light emission discussed in Sec. 5 is, of course, quite analogous to the nuclear disintegration process considered here. Equation (5.1), expressed in terms of spherical harmonics, has the form

$$I(\theta)\rho\propto\frac{10}{9}\left\{[1-\tfrac{1}{10}P_{20}(\theta)]\mathbb{1}+\frac{6^{\frac{1}{2}}}{10}P_{22}(\theta)\sigma_\zeta\right\},\quad(11.2)$$

with $P_{20}=(3/2)\cos^2\theta-(\frac{1}{2})$ and $P_{22}=(3/8)^{\frac{1}{2}}\sin^2\theta$. In the calculation of (11.2) the coefficients of P_{20} and P_{22} have been reduced by a factor $\frac{1}{10}$ owing to a $\bar{W}$ coefficient which takes into account the quantum number $j'=2$ of the atom after the radiation emission. If the atom were left in an s state ($j'=0$), this factor $\frac{1}{10}$ would not appear; indeed the light emitted in a transition from ($l=1, m=0$) to $l=0$ is fully polarized and its angular distribution is well modulated, as indicated by

$$I(\theta)\rho\propto[1-P_{20}(\theta)]\mathbb{1}+6^{\frac{1}{2}}P_{22}(\theta)\sigma_\zeta=\tfrac{3}{2}\sin^2\theta(\mathbb{1}+\sigma_\zeta).\quad(11.3)$$

(c) Interaction Representation—Feynman Calculus

Equation (3.18) shows that the calculations of a mean value $\langle Q\rangle$ in the Heisenberg representation ($Q=Q(t)$, $\rho=\mathrm{const}$) and in the Schroedinger representation ($Q=\mathrm{const}$, $\rho=\rho(t)$) are manifestly equivalent. When the Hamiltonian consists of two or more terms, as in (10.2), other representations are possible, in which the time variations induced by different terms of the Hamiltonian are separated out. The Hamiltonian (10.2) contains independent-system terms H_a+H_b and a coupling term V. In the "interaction representation" the variations induced by H_a+H_b may be applied to Q and the residual influence of V applied to ρ. This representation is particularly convenient when $Q=Q^{(a)}$ and commutes with H_b, because H_b causes then no variation of Q.

The formal derivation and the applications of the interaction representation are facilitated by Feynman's operator calculus (F51) which shows how to "disentangle" operators contained in the same exponential function, to find

$$\begin{aligned}\exp[-i\hbar^{-1}(H_a+H_b+V)t]&\\ &=\exp[-i\hbar^{-1}(H_a+H_b)t]\\ &\quad\times\exp\left[-i\hbar^{-1}\int_0^t V(t')dt'\right],\quad(11.4)\end{aligned}$$

with

$$\begin{aligned}V(t)=\exp[i\hbar^{-1}(H_a+H_b)t]V&\\ &\times\exp[-i\hbar^{-1}(H_a+H_b)t],\quad(11.5)\end{aligned}$$

and the reciprocal formula

$$\begin{aligned}&\exp[i\hbar^{-1}(H_a+H_b+V)t]\\ &\quad=\exp\left[i\hbar^{-1}\int_0^t V(t')dt'\right]\exp[i\hbar^{-1}(H_a+H_b)t].\quad(11.6)\end{aligned}$$

Equation (10.6) becomes, in this representation,

$$\begin{aligned}\langle Q^{(a)}\rangle=\mathrm{Tr}_{ab}\bigg\{Q^{(a)}(t)\exp\left[-i\hbar^{-1}\int_0^t V(t')dt'\right](\rho^{(ab)})_0&\\ \times\exp\left[i\hbar^{-1}\int_0^t V(t')dt'\right]\bigg\},\quad(11.7)\end{aligned}$$

where $Q^{(a)}(t)$ is independent of H_b, and (10.1) becomes

$$(\rho^{(a)})_t=\exp(-i\hbar^{-1}H_a t)$$
$$\times\mathrm{Tr}_b\left\{\exp\left[-i\hbar^{-1}\int_0^t V(t')dt'\right](\rho^{(ab)})_0\right.$$
$$\left.\times\exp\left[i\hbar\int_0^t V(t')dt'\right]\right\}\exp(i\hbar^{-1}H_a t), \quad (11.8)$$

where H_b has dropped out from the expression of the trace.

Feynman (F51) has also pointed out how, given an operator, such as V, acting on both systems a and b, one may calculate its matrix elements between states of b only, so that these matrix elements are still operators with regard to a. To this end one may expand the operator V into operators U_i of a, according to (6.2), $V=\sum_i V_i U_i$, the coefficients V_i now being operators of b only. We have then

$$V(t)=\textstyle\sum_i V_i(t)U_i(t)=\sum_i \exp(i\hbar^{-1}H_b t)V_i$$
$$\times\exp(-i\hbar^{-1}H_b t)\exp(i\hbar^{-1}H_a t)U_i$$
$$\times\exp(-i\hbar^{-1}H_a t), \quad (11.9)$$

which enables us to calculate directly the matrix elements of $V_i(t)$ only, or also expressions such as the $\mathrm{Tr}_b[\cdots]$ in (11.8).

In the calculation pertaining to b one may treat the operators $U_i(t)$ as though they were ordinary functions rather than operators. A main point of Feynman's method lies in utilizing variable indices, such as t, to keep track of the correct ordering of the operators of the system a, even though the usual operator algebra rules may have been formally disregarded in the calculation relating to b.

(d) Perturbation Expansions

The results of the usual time-dependent perturbation theory are easily obtained in a more general form by utilizing the expansion (F51)

$$\exp\left[\int_0^t A(t')dt'\right]=1+\int_0^t A(t')dt'$$
$$+\int_0^t dt'\int_0^{t'} dt''A(t')A(t'')+\cdots, \quad (11.10)$$

which enables one to substitute in (11.7) or (11.8),

$$\exp\left[-i\hbar^{-1}\int_0^t V(t')dt'\right](\rho^{(ab)})_0\exp\left[i\hbar^{-1}\int_0^t V(t')dt'\right]$$
$$=(\rho^{(ab)})_0-i\hbar^{-1}\int_0^t dt'[V(t'),(\rho^{(ab)})_0]$$
$$-\hbar^{-2}\int_0 dt'\int_0^{t'} dt''[V(t'),[V(t''),(\rho^{(ab)})_0]]+\cdots. \quad (11.11)$$

The usual perturbation formulas are obtained by representing (11.11) in the scheme of eigenstates of H_A+H_B; the matrix elements of $V(t)$ are then simple exponentials in t, as in (3.15), and the integrations over t' and t'' may be carried out analytically.

(e) Correlation Functions

When the systems a and b are initially independent, i.e. when $(\rho^{(ab)})_0=(\rho^{(a)})_0(\rho^{(b)})_0$, V is expanded in the form (11.9), and the perturbation expansion (11.11) is utilized, the $\mathrm{Tr}_b[\cdots]$ in (11.7) or (11.8) depends on b through a series of terms of the form

$$\mathrm{Tr}\{V_i(t')V_j(t'')\cdots V_l(t^{(n)})(\rho^{(b)})_0\}$$
$$=\langle V_i(t')V_j(t'')\cdots V_l(t^{(n)})\rangle. \quad (11.12)$$

These parameters are time-correlation functions of a kind often considered in statistical mechanics, that is, mean values of products of physical quantities taken at different instants of time. (However, in classical mechanics the averaging is usually done by an integration over time intervals, whereas in (11.12) we deal with an ensemble average.)

In the frequent case where b is initially in a steady state, i.e., when $(\rho^{(b)})_0$ commutes with H_b, the correlation functions (11.12) depend only on the time intervals $t'-t''$, $t''-t'''$, etc. but not on the absolute values of t', $t''\cdots$. In this event $\langle V_i(t')\rangle$ is altogether independent of t', and the pair correlation function $\langle V_i(t')V_j(t'')\rangle$ depends only on $t'-t''$ and may be written as $\langle V_iV_j(t''-t')\rangle$.

If the expansion $V=\sum_i U_iV_i$ is not utilized, we have instead of (11.12) expressions of the form $\mathrm{Tr}_b\{V(t')V(t'')\cdots V(t^{(n)})(\rho^{(b)})_0\}$ which are operators of the system a, whereas the parameters (11.12) are numerical coefficients of the operators $U_i(t')U_j(t'')\cdots$. If $(\rho^{(b)})_0$ commutes with H_b, the dependence on the absolute time scale can still be brought out of the correlation function; we have then

$$\mathrm{Tr}\{V(t')V(t'')\cdots V(t^{(n)})(\rho^{(b)})_0\}$$
$$=\exp(i\hbar^{-1}H_a t')\ \mathrm{Tr}_b\{VV(t''-t')\cdots V(t^{(n)}-t')(\rho^{(b)})_0\}$$
$$\cdot\exp(-i\hbar^{-1}H_a t'). \quad (11.13)$$

The correlation functions are proving increasingly useful for representing and classifying the properties of a system which are relevant to its interaction with other systems. For a system initially in a steady state, the static parameters $\langle V_i\rangle$ represent constant fields that produce elastic scattering effects, whereas the pair correlations $\langle V_i(0)V_j(t)\rangle$ represent to first-order inelastic interaction parameters. For example, Van Hove (VH54) has represented the neutron scattering properties of a crystal in terms of a pair correlation between the positions of nuclei $\mathbf{r}_i$, $\mathbf{r}_j$ in the crystal

$$G(\mathbf{r},t)=N^{-1}\left(\sum_{i,j=1}^{N}\int d\mathbf{r}'\delta(\mathbf{r}+\mathbf{r}_i(0)-\mathbf{r}')\right.$$
$$\left.\times\delta(\mathbf{r}'-\mathbf{r}_j(t))\right). \quad (11.14)$$

The function $G(\mathbf{r},t)$ is a macroscopic property of the crystal, but its definition is quantum-mechanical, atomistic, rather than phenomenologic. As a further example, the dielectric constant of a material, determined by electron polarizability, is defined, in terms of the $G(\mathbf{r},t)$ function for electrons (F56), by

$$\epsilon(\omega)=\left\{1+\frac{Ne^2}{\hbar}\int_0^\infty dte^{-i\omega t}\int d\mathbf{r}\right.\\ \left.\times\int d\mathbf{r}'\frac{G(\mathbf{r}',t)-G(\mathbf{r}',t)^*}{|\mathbf{r}-\mathbf{r}'|}\right\}^{-1}. \quad (11.15)$$

(f) Probability of Energy or Momentum Transfers

There is another approach which leads one to express the probability of an interaction effect in terms of a correlation function (AP47). Suppose one inquires about the probability $P(\omega)d\omega$ that system a transfers to b an energy amount between $\hbar\omega$ and $\hbar(\omega+d\omega)$ in the time interval from 0 to t, irrespective of other effects of the interaction. This probability is represented by the mean value of an operator $P(\omega)$ which is diagonal in the scheme of energy eigenstates of a and has in this scheme the form $\delta(E_{at}-E_{a0}+\hbar\omega)$, where E_{at} and E_{a0} are eigenvalues of H_a pertaining, respectively, to states at the times t and 0. The mean value of P is to be evaluated according to (10.6). We represent $P(\omega)$ in the form

$$P(\omega)=\delta(E_{at}-E_{a0}+\hbar\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty}d\tau\exp(i\omega\tau)\\ \times\exp(i\hbar^{-1}E_{at}\tau)\exp(-i\hbar^{-1}E_{a0}\tau) \quad (11.16)$$

and notice that both E_{at} and E_{a0} may be represented by the Hamiltonian H_a operating, respectively, after and before the interaction. Entering into (10.6) $Q^{(a)}=P(\omega)$, we may write, in operator notation independent of a specific matrix representation,

$$\langle P(\omega)\rangle=\frac{1}{2\pi}\int_{-\infty}^{\infty}d\tau\exp(i\omega\tau)\\ \times\mathrm{Tr}_{ab}\{\exp(i\hbar^{-1}H_a\tau)\exp(-i\hbar^{-1}Ht)\\ \times\exp(-i\hbar^{-1}H_a\tau)(\rho^{(ab)})\exp(i\hbar^{-1}Ht)\}. \quad (11.17)$$

Further, utilizing the interaction representation (11.7) and the expansion (11.9), and considering that for any operator $F^{(a)}(t)$ in the interaction representation $\exp(i\hbar^{-1}H_a\tau)F^{(a)}(t)\exp(-i\hbar^{-1}H_a\tau)=F^{(a)}(t+\tau)$, we have

$$\langle P(\omega)\rangle=\frac{1}{2\pi}\int_{-\infty}^{\infty}d\tau\exp(i\omega\tau)\\ \times\mathrm{Tr}\left\{\exp\left[-i\hbar^{-1}\int_0^t\textstyle\sum_i U_i(t'+\tau)V_i(t')\right](\rho^{(ab)})_0\right.\\ \left.\times\exp\left[i\hbar^{-1}\int_0^t\textstyle\sum_j U_j(t')V_j(t')\right]\right\}. \quad (11.18)$$

In the first Born approximation the integrations over t' are carried out through the second order term of the perturbation expansion (11.11), and (11.18) becomes

$$\langle P^{(a)}(\omega)\rangle=\frac{t}{\hbar}\int_{-\infty}^{\infty}d\tau\exp(i\omega\tau)\,\mathrm{Tr}_{ab}\{[\textstyle\sum_i\bar{U}_i(\tau)\bar{V}_i]\\ \times(\rho^{(ab)})_0[\textstyle\sum_j\bar{U}_j\bar{V}_j]\}, \quad (11.19)$$

where $\bar{U}_i\bar{V}_i$ represents that portion of the operator product U_iV_i which is diagonal in the total energy and induces transitions with energy conservation, namely

$$\bar{U}_i\bar{V}_i=(2\pi\hbar)^{-1}\int_{-\infty}^{\infty}dtU_i(t)V_i(t). \quad (11.20)$$

It follows that $\bar{U}_i(\tau)\bar{V}_i=\bar{U}_i\bar{V}_i(-\tau)$.

Expressions like (11.19) and (11.20) show explicitly how transition probabilities depend on a time variable whose role is obscured in the usual quantum-mechanical formulas. For example, the absorption or emission of light by an atomic electron depends on the correlation between the electron's position $\mathbf{r}$ or velocity $\mathbf{v}$ at time intervals τ. (In this problem, we have $\sum_i U_iV_i=e\mathbf{v}\cdot\mathbf{A}(\mathbf{r})$, where $\mathbf{A}$ is the vector potential.) The probability of emitting or absorbing a photon $\hbar\omega$ is proportional to the spectral component ω of the correlation function. Van Hove was led to consider the correlation function G of (11.14) (VH54) through a formula similar to (11.19). Wick (W54a) utilized such a formula as the point of departure of an approximate calculation of neutron scattering in which the correlation function is expanded into powers of τ, to reflect the fact that a neutron scattering process is short as compared to the vibration period of a crystal lattice.

The probability of a momentum transfer $\hbar\mathbf{q}$ may be treated much like that of an energy transfer, in terms of an operator $\delta(\mathbf{p}_{at}-\mathbf{p}_{a0}+\hbar\mathbf{q})$ expanded into a triple Fourier integral. The operators $\exp(\pm i\hbar^{-1}\mathbf{p}_a\cdot\boldsymbol{\xi})$, which replace $\exp(\pm i\hbar^{-1}H_a\tau)$ in (11.17) have the effect of shifting all space coordinates in the operator U_i by $\boldsymbol{\xi}$, so that the trace takes the form of a space correlation function of the type $\langle f(\mathbf{r})f(\mathbf{r}+\boldsymbol{\xi})\rangle$.

(g) Wangsness-Bloch Equation for Irreversible Processes

Wangsness and Bloch (WB53) treated the relaxation of nuclear spin orientation, in essence, along the lines indicated at the end of Sec. 10. The variable orientation of the spin of a fixed atomic nucleus constitutes the system a, whose Hamiltonian H_a represents the coupling with an external magnetic field $\mathbf{H}$, and the system of all surrounding particles constitutes b, the Hamiltonian H_b being unspecified. At the initial time $t=0$, the states of the two systems are assumed to be uncorrelated, so that $(\rho^{(ab)})_0=(\rho^{(a)})_0(\rho^{(b)})_0$, according to (3.20), with $(\rho^{(a)})_0$ unspecified and $(\rho^{(b)})_0$ assumed to be a state of thermal equilibrium according to (4.17). The coupling V of a and

b is also left unspecified except that $\mathrm{Tr}_b V$ is assumed to vanish, though only for simplicity.

The state of spin orientation at a somewhat later time $t=\tau$, represented by $(\rho^{(a)})_\tau$, is then given by (11.8) but τ is assumed to be sufficiently small to warrant application of the perturbation expansion (11.11) extended to second order terms only. Disregarding further some minor side effects, one finds

$$(\rho^{(a)})_\tau=\exp(-i\hbar^{-1}H_a\tau)\,\mathrm{Tr}_b\left\{(\rho^{(a)})_0(\rho^{\ b)})_0\right.$$
$$\left.-\frac{\pi\tau}{\hbar}[\bar{V},[\bar{V},(\rho^{(a)})_0(\rho^{(b)})_0]]\right\}\exp(i\hbar^{-1}H_a\tau),\qquad(11.21)$$

where $\bar{V}$ represents that portion of the operator V which is diagonal in the total energy and is defined according to (11.20). (The first-order term is missing because $\mathrm{Tr}_b V$ is assumed to vanish.)

To extend the calculation to larger intervals $t>\tau$ one would need, in general, to know not only $(\rho^{(a)})_\tau$ but the density matrix of the whole system $(\rho^{(ab)})_\tau$. At this point Wangsness and Bloch take into account that b is a large system through which the effects of interaction with the spin a dissipate away quickly. Therefore, it is argued, the state of b will not have changed significantly during τ. The actual density matrix $(\rho^{(ab)})_\tau$, resulting from $(\rho^{(a)})_0(\rho^{(b)})_0$ by reversible transformation, is replaced with

$$(\rho_s^{(ab)})_\tau=(\rho^{(a)})_\tau(\rho^{(b)})_0.\qquad(11.22)$$

The evolution in time of the combined system becomes thereby irreversible.

If this $(\rho_s^{(ab)})_\tau$ is entered into (11.21) in the place of $(\rho^{(a)})_0(\rho^{(b)})_0$, the result is $(\rho^{(a)})_{2\tau}$. The operation can be repeated chainwise yielding $(\rho^{(a)})_t$ at successive time intervals τ, 2τ, 3τ, $\cdots n\tau\cdots$. This repetitive operation is analogous to the numerical integration of a differential equation. Indeed, the variations of $(\rho^{(a)})_t$ thus obtained obey a differential equation, where the interval τ is treated formally as a differential dt[15]

$$\frac{d\rho^{(a)}}{dt}=-\frac{i}{\hbar}[H_a,\rho^{(a)}]-\frac{\pi}{\hbar}\mathrm{Tr}_b[\bar{V},[\bar{V},\rho^{(a)}(\rho^{(b)})_0]].\qquad(11.23)$$

Wangsness and Bloch (WB) pointed out that (11.23) is a quantum analog of the Boltzmann equation of statistical mechanics, and generalizes the Schroedinger equation to incorporate irreversible transformations. Expansion of $\rho^{(a)}$ into orthogonal operators, $\rho^{(a)}=\sum_i\rho_iU_i$, brings the WB equation to a form analogous to (6.9),

$$d\rho_i/dt=\sum_k(\Omega_{ik}+M_{ik})\rho_k,\qquad(11.24)$$

with Ω_{ik} as in (6.11) and

$$M_{ik}=-\frac{\pi}{\hbar}\mathrm{Tr}_{ab}\{U_i[\bar{V},[\bar{V},U_k(\rho^{(b)})_0]]\}.\qquad(11.25)$$

Equation (11.24) represents a transformation of the vector $\boldsymbol{\rho}^{(a)}$ in the Hilbert space of Sec. 7, but we deal no longer with a simple rotation because M_{ik} is not antisymmetric. With $U_i=U_0\propto\mathbb{1}$, we have $M_{0k}=\Omega_{0k}=0$ (but $M_{i0}\neq0$ in general), so that $\partial\rho_0/\partial t=0$ as required; it also follows that the determinant of $\Omega_{ik}+M_{ik}$ vanishes, so that (11.24) has a steady state solution.

Further expansion $\bar{V}=\sum_r U_r\bar{V}_r$ makes it possible to represent the matrix M_{ik} in terms of separate properties of the systems a and b. The relevant properties of a depend only on its mechanical structure as represented by relationships among the operators U_i. They depend in particular on the coefficients $c_{ik}{}^p$ of (6.10) and on the corresponding symmetric coefficients.

$$s_{ik}{}^p=\mathrm{Tr}[(U_iU_k+U_kU_i)U_p].\qquad(11.26)$$

Entering the expansion of $\bar{V}$ into (11.25), working out the double commutator, and separating out symmetric and antisymmetric terms, one finds

$$M_{ik}=-\sum_{rs}\left\{A_{ikrs}\frac{\pi}{2\hbar}\langle\bar{V}_r\bar{V}_s+\bar{V}_s\bar{V}_r\rangle\right.$$
$$\left.+B_{ikrs}\frac{\pi}{2\hbar}\langle i[\bar{V}_r,\bar{V}_s]\rangle\right\}\qquad(11.27)$$

where

$$A_{ikrs}=A_{kisr}=\sum_p c_{ir}{}^pc_{ks}{}^p,\qquad(11.28)$$

$$B_{ikrs}=-B_{iksr}=\sum_p c_{ir}{}^ps_{ks}{}^p$$
$$=\sum_p(\tfrac{1}{2}s_{ik}{}^pc_{rs}{}^p+\tfrac{1}{4}c_{ir}{}^ps_{ks}{}^p-\tfrac{1}{4}c_{kr}{}^ps_{is}{}^p$$
$$-\tfrac{1}{4}c_{is}{}^ps_{kr}{}^p+\tfrac{1}{4}c_{ks}{}^ps_{ir}{}^p).\qquad(11.29)$$

The coefficients A and B represent the properties of a. They obey selection rules, arising from the invariance of traces under coordinate rotation or under reflection in space or time. For example, the $c_{ik}{}^p$ vanish unless the product of the parities of U_i, U_k, and U_p is odd, and the $s_{ik}{}^p$ vanish unless it is even, so that $A_{ikrs}=0$ if the product $U_iU_kU_rU_s$ is odd and $B_{ikrs}=0$ when it is even.

The mean values $\langle\bar{V}_r\bar{V}_s+\bar{V}_s\bar{V}_r\rangle$ and $\langle i[\bar{V}_r,\bar{V}_s]\rangle$ represent the relevant properties of b. They are essentially equivalent to the admittance matrix, with rows and columns indicated by r and s, which has been introduced by Callen and co-workers (CW51, CBJ52). These authors considered the reaction of a large system, like b, to an external time variable potential $Q_s\exp(i\omega t)$, representing the reaction by the mean value of the rate of change of another operator $\dot{Q}_r$, as calculated by first-order perturbation. The Feynman calculus exploits the analogy between the reaction of b to an external po-

---

[15] In the derivation of (11.21) it is assumed that τ is not infinitesimal but, rather, that is sufficiently large for the spectrum of the interaction, $\int_{-\infty}^{\infty}V(t)\exp(-i\omega t)dt$ to be effectively constant over an interval $\delta\omega=\tau^{-1}$. Therefore (11.21) and (11.23) represent correctly only semimacroscopic time variations of $\rho^{(a)}$, leaving out variations on a finer scale of time intervals $\lesssim(\delta\omega)^{-1}$.

tential and to the interaction V with another quantum-mechanical system a, by permitting one to calculate the reaction to a without considering at every step that V is also an operator of a.[16]

A system a with N orthogonal eigenstates has N^2 operators U_i, so that the matrix $\Omega_{ik}+M_{ik}$ has N^2 rows and columns. The relevant characteristics of b are represented by N^4 coefficients $\langle \bar{V}_r\bar{V}_s+\bar{V}_s\bar{V}_r\rangle$ and as many $\langle i[\bar{V}_r,\bar{V}_s]\rangle$. The matrix $\Omega_{ik}+M_{ik}$ has N^2 complex eigenvalues, whose real parts are time-relaxation constants; one of these eigenvalues vanishes, in correspondence to the steady states, the remaining N^2-1 relaxation constants determine the rate of approach to the steady state. All these parameters need not be different and their number is often greatly reduced owing to symmetry.

In the nuclear induction with spin $\frac{1}{2}$, there could be $N^2-1=3$ relaxation constants, but axial symmetry about the main magnetic field reduces this number to 2, corresponding to the well-known longitudinal and transverse relaxation times. For spin 1 we have $N^2-1=8$ constants, reduced to 5 by axial symmetry. However Wangsness and Bloch argue that dipole and quadrupole polarizations are effectively uncoupled in this case.[17] Hence the magnetic dipole polarization, which is responsible for nuclear induction, behaves for $j=1$ as though j were $\frac{1}{2}$.

(h) Irreversibility Approximation by Expansion in Cumulants

Instead of introducing irreversibility through the convention (11.22), one may formulate a systematic approximation procedure (F54b), which, when applied to first order, reduces the exact expression (11.8) for $(\rho^a)_t$ to a solution of the Wangsness-Bloch Eq. (11.23). Qualitatively, if a system b is large and quickly dissipates the effect of disturbances, one expects b to show a "lack of memory" in its interactions with other systems, such that the time variations of an operator $V(t)$ in the neighborhood of time instants t' and t'' would be uncorrelated, for $t'-t''$ sufficiently large. Mathematically, this lack of correlation should cause the mean value of an operator such as $\langle V_i(t')V_j(t'')\cdots V_l(t^n)\rangle$ in (11.12) to approach rapidly the product of mean values $\langle V_i\rangle\langle V_j\rangle\langle V_l\rangle$ as the time intervals $t'-t''\cdots t'-t^{(n)}$ increase.

In order to take advantage of this circumstance one considers in particular the departures of an operator from its mean value, defining for example $\mathcal{V}(t)=V(t)-\langle V(t)\rangle_b$ (where $\langle A\rangle_b$ means Tr_bA), so that $\langle\mathcal{V}(t)\rangle_b=0$

$$V(t)=\langle V(t)\rangle_b+\mathcal{V}(t) \tag{11.30}$$

and

$$\langle VV(t)\rangle_b=\langle V\rangle_b\langle V(t)\rangle_b+\langle\mathcal{V}\mathcal{V}(t)\rangle_b. \tag{11.31}$$

The mean value $\langle\mathcal{V}\mathcal{V}(t)\rangle_b$, is the correlation of "deviates." It represents the correlation effect proper and is expected to vanish when t increases above an effective memory range, or relaxation time, of b. This procedure may be continued by separating out deviates of successively higher orders, defining, for example,

$$\mathcal{V}(t')\mathcal{V}(t'')=\langle\mathcal{V}(t')\mathcal{V}(t'')\rangle_b+\mathcal{W}(t')\mathcal{W}(t''), \tag{11.32}$$

so that the deviates $\mathcal{W}$ have not only $\langle\mathcal{W}(t)\rangle_b=0$ but also $\langle\mathcal{W}(t')\mathcal{W}(t'')\rangle_b=0$; the $\mathcal{W}$'s contribute only to quadruple and higher correlations.

When expressions like (11.30) and (11.32) are entered in the perturbation expansion (11.11), the mean values of different orders $\langle V(t)\rangle_b$, $\langle\mathcal{V}(t')\mathcal{V}(t'')\rangle_b$, etc. can be sorted out from the expansion. Leaving out the complications due to commutation with $(\rho^{(ab)})_0$ (see F54b), and using (11.13) the reader may verify that the simpler expansion

$$\begin{aligned}\mathrm{Tr}\left\{\exp\left(-i\hbar^{-1}\int_0^t V(t')dt'\right)\rho^{(b)}\right\}\\=\mathrm{Tr}_b\left\{\left[1-i\hbar^{-1}\int_0^t V(t')dt'\right.\right.\\\left.\left.-\hbar^{-2}\int_0^t dt'\int_0^{t'} dt''V(t')V(t'')+\cdots\right]\rho^{(b)}\right\}\end{aligned} \tag{11.33}$$

condenses to the exponential form

$$\begin{aligned}\exp\left\{-i\hbar^{-1}\int_0^t dt'\exp(i\hbar^{-1}H_at')\right.\\\times\left[\langle V\rangle_b-i\hbar^{-1}\int_0^{t'}dt''\langle\mathcal{V}\mathcal{V}(t''-t')\rangle_b\right.\\\left.-\hbar^{-2}\int_0^{t'}dt''\int_0^{t''}dt'''\langle\mathcal{V}\mathcal{V}(t''-t')\mathcal{V}(t'''-t')\rangle_b+\cdots\right]\\\left.\times\exp(-i\hbar^{-1}H_at')\right\}.\end{aligned} \tag{11.34}$$

The average of the exponential operator in (11.33) has been thereby reduced to the exponential function of a sequence of averaged operators, which are correlation functions of increasing order. The reduction of (11.33) to (11.34) is analogous to the equation of probability theory $\langle\exp(t\xi)\rangle=\exp(\sum_\nu t^\nu\kappa_\nu/\nu!)$, where ξ is a random variable and the κ_ν are the cumulants $\langle\xi\rangle$, $\langle(\xi-\langle\xi\rangle)^2\rangle$, $\langle(\xi-\langle\xi\rangle)^3\rangle\cdots$ of its distribution (C46).

---

[16] The analog of Callen's time variation $\exp(i\omega t)$ of the external disturbance is included in the WB problem through the definition

$$\begin{aligned}V_s(t)=\mathrm{Tr}_a\{U_sV(t)\}=\mathrm{Tr}\{\exp(i\hbar^{-1}H_at)U_s\exp(-i\hbar^{-1}H_at)\\\times\exp(i\hbar^{-1}H_bt)V\exp(-i\hbar^{-1}H_bt)\}\end{aligned}$$

which contains a factor depending on all eigenfrequencies of H_a.

[17] The coefficients $\langle\bar{V}_r\bar{V}_s+\bar{V}_s\bar{V}_r\rangle$ and $\langle i[\bar{V}_r,\bar{V}_s]\rangle$ can be classified according to the multipolarity of the operators U_r and U_s. If U_r is a dipole and U_s a quadrupole operator, $\bar{V}_r$ represents a magnetic field and $\bar{V}_s$ an electric field, and WB assume that the fluctuations of these different fields are uncorrelated so that the mixed-field coefficients vanish. It follows then from the parity selection rules of the coefficients A_{ikrs} and B_{ikrs} that the parameters ρ_i and ρ_k pertaining to magnetic and electric polarizations will remain uncoupled.

The expansion (11.34) affords a suitable point of departure for an irreversibility approximation. Three steps are involved in deriving from (11.34) a simpler expression equivalent to the Wangsness-Bloch Eq. (11.23): (1) It must be assumed that, because of the internal mechanics of system b, the correlation functions in (11.34) vanish rapidly as $t'-t''$, $t'-t'''\cdots$ increase, so that, as t' increases, the integrals in the brackets reach rapidly a limiting value corresponding to $t'=\infty$ and thus are no longer explicit functions of t' (short-memory approximation). (2) The whole expression in the brackets can be handled as a single operator of the system a, acting at the time t', even though some portions of it actually should operate at times t'', $t'''\cdots$ somewhat earlier than t'. Disregarding these fine distinctions in the time scale is equivalent to treating the time intervals τ as differentials in the derivation of (11.23). (3) The terms of third and higher order can be disregarded. This assumption relates to the short-memory approximation (1) because high-order interaction effects would build up if allowed to accumulate for a long time. It implies that second-order perturbation treatment is adequate, as in (11.21), to describe the effects of interaction upon a during each memory interval.

Thus the whole effect of irreversible interaction with system b is represented by pair correlation integrals of the type $\int_0^\infty dt \langle \mathcal{V}\mathcal{V}(-t)\rangle_b$. Upon expansion into orthogonal operators $\mathcal{V}(t)=\sum_s U_s\mathcal{V}_s(t)$ these integrals depend on b through mean values $\langle \mathcal{V}_r\mathcal{V}_s(-t)\rangle$ which are again equivalent to Callen's admittance matrix mentioned at the end of Sec. 11g.

REFERENCES

AP47 A. Akhiezer and I. Pomeranchuk, J. Phys. (U.S.S.R.) **11**, 167 (1947).
B46 F. Bloch, Phys. Rev. **70**, 460 (1946).
B56 F. Bopp, Z. Physik **144**, 13 (1956).
BR45 F. Bloch and I. I. Rabi, Revs. Modern Phys. **17**, 237 (1945).
C33 H. B. G. Casimir, Helv. Phys. Acta **6**, 287 (1933).
C46 H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, 1946), pp. 185–186.
CBJ52 Callen, Barasch, and Jackson, Phys. Rev. **88**, 1382 (1952).
CJ53 F. Coester and J. Jauch, Helv. Phys. Acta **26**, 3 (1953).
CW51 H. B. Callen and T. A. Welton, Phys. Rev. **83**, 34 (1951).
D52 R. H. Dalitz, Proc. Phys. Soc. (London) **A65**, 175 (1952).
E37 W. M. Elsasser, Phys. Rev. **52**, 987 (1937).
F49 U. Fano, J. Opt. Soc. Am. **39**, 859 (1949).
F51 R. Feynman, Phys. Rev. **84**, 108 (1951).
F51a U. Fano, NBS Report 1214, unpublished (1951).
F53 U. Fano, Phys. Rev. **90**, 577 (1953).
F54a U. Fano, Phys. Rev. **93**, 121 (1954).
F54b U. Fano, Phys. Rev. **96**, 869 (1954).
F56 U. Fano, Phys. Rev. **103**, 1202 (1956).
FMcD51 D. L. Falkoff and J. E. McDonald, J. Opt. Soc. Am. **41**, 862 (1951).
FR57 U. Fano and G. Racah, *Irreducible Tensors* (Academic Press, Inc., New York).
H54 W. Heitler, *The Quantum Theory of Radiation* (Clarendon Press, Oxford, 1954), third edition.
IZ51 J. Irving and R. Zwanzig, J. Chem. Phys. **19**, 1173 (1951).
K55 T. Kotani, Progr. Theoret. Phys. **14**, 379 (1955).
L50 S. P. Lloyd, Phys. Rev. **80**, 118 (1950).
L51 B. A. Lippmann, Phys. Rev. **81**, 162 (1951).
M48 J. E. Moyal, Proc. Cambridge Phil. Soc. **45**, 100 (1948).
N27 J. von Neumann, Göttinger Nachr. 245 and 273 (1927).
N55 J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, 1955).
S50 J. A. Spiers, Phys. Rev. **80**, 491 (1950).
S52 G. G. Stokes, Proc. Cambridge Phil. Soc. **9**, 399 (1852).
S55 R. L. Stratonovich, J. Exptl. Theoret. Phys. U.S.S.R. **28**, 547 (1955).
T38 R. C. Tolman, *The Principles of Statistical Mechanics* (Clarendon Press, Oxford, 1938).
TG49 H. A. Tolhoek and S. R. de Groot, Physica **15**, 833 (1949).
TG51 H. A. Tolhoek and S. R. de Groot, Phys. Rev. **83**, 189 (1951).
VH54 L. Van Hove, Phys. Rev. **95**, 249 (1954).
VH55 L. Van Hove, Physica **21**, 517 (1955).
W31 H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover Publications, New York, 1931).
W32 E. Wigner, Phys. Rev. **40**, 749 (1932).
W53 L. Wolfenstein, Phys. Rev. **92**, 123 (1953).
W54a G. C. Wick, Phys. Rev. **94**, 1228 (1954).
W54b L. Wolfenstein, Phys. Rev. **96**, 164 (1954).
WA52 L. Wolfenstein and J. Ashkin, Phys. Rev. **85**, 947 (1952).
WWW52 Wick, Wightman, and Wigner, Phys. Rev. **88**, 101 (1952).