

CN Chapter 1

CT Classical or Quantum

Probabilities. Or both?

The description of nature through quantum theory is intrinsically probabilistic, and statistical methods are developed in order to get information of physical observables. This distinctive feature has its roots in the probabilistic interpretation of the *wave function*, in conjunction with the Uncertainty Principle. As discussed in standard textbooks, quantum probabilities embody the dual nature of matter, reproducing phenomena such as interference and diffraction. This statistical representation, even for a one particle system, requires an *ensemble* of similarly prepared systems, in order to generate a set of well defined probabilities for all quantities. The ensemble is a conceptual set of an ‘infinite’ number of replicas of the same system that do not coexist in space or time.

CHAPTER 1 — MANUSCRIPT

When all systems of the ensemble are represented by the same *wave function* or *state vector* $|\psi\rangle$, we say that we have prepared a *pure state*. In Statistical Mechanics, $|\psi\rangle$ represents a possible *microstate* of the system. Physical predictions of an observable quantity \mathbf{A} are given in terms of averages or mean values of the type:

$$\langle \mathbf{A} \rangle = \langle \psi | \mathbf{A} | \psi \rangle, \quad (1.1)$$

where the bracket is calculated following well established quantum rules. In turn, statistical fluctuations of quantum origin are obtained through:

$$\langle \mathbf{A}^2 \rangle - \langle \mathbf{A} \rangle^2 = \langle \psi | (\mathbf{A} - \langle \mathbf{A} \rangle)^2 | \psi \rangle, \quad (1.2)$$

where we have assumed that the state ψ is normalized, *i.e.* $\langle \psi | \psi \rangle = 1$. The time evolution of this microstate is dictated by the Schrödinger equation, as long as no measurement is made on the system. We will not pursue here a thorough discussion on the fundamental concepts of Quantum Mechanics, and will assume that the reader has the proper background to follow this set of lectures [1, 2]. According to Quantum Mechanics, a pure state is the maximal information we can have of a quantum system, and represents an exception rather than the rule. In most situations, we do not know the wave function with certainty. A typical example is given by a system of coordinate \mathbf{x} which is coupled to another system of coordinate \mathbf{y} (the latter may be a thermal bath that keeps the system at a constant temperature). In general, we cannot assign a wave function to our system \mathbf{x} , which is a part of the whole system (\mathbf{x}, \mathbf{y}) . In

CHAPTER 1 — MANUSCRIPT

this case, we say that the system \mathbf{x} is in a *mixed state* and we have to look for a more general mathematical object to describe this situation. This lack of information closely resembles the classical statistical problem. In fact, in many instances this case is denominated as *incoherent mixture*, to mean the lack of quantum coherence and the absence of interference effects. As will be discussed later on, quantum and thermal fluctuations compete in this scenario as function of temperature. Quantum coherence is realized at low temperatures, and asymptotically at zero temperature the system is driven exclusively by quantum fluctuations. As long as the temperature is increased, the system develops thermal fluctuations, which eventually dominate the statistics, leading to decoherence of quantum effects (classical statistics). The description of the general ensemble, encompassing all cases, is attained quite naturally employing the so called *state operator* ρ (or *density operator*), which is a generalization of the wave function concept. The operator ρ is the relevant quantity to construct quantum statistical mechanics. The name *density* is reminiscent of the classical function $\rho(q, p)$ that yields the density of points in phase space to perform classical statistics. A single point in phase space represents one of the systems of the ensemble. An important task to be accomplished is to relate the quantum density operator ρ with the classical function $\rho(q, p)$. In the next section, we develop the basis of the mathematical formalism.

A

1.1 Ensembles and Density Operator

We begin by defining the different ensembles we encounter to describe a physical system [3].

Definition 1 *Pure Ensemble.* *It is a collection or set of identical physical systems, such that all members of the ensemble (systems) are characterized by the same state vector or ket $|\psi\rangle$. This is the usual case we encounter in standard quantum mechanics textbooks, when the state of a system is represented by a single wave function. This ensemble represents a microstate.*

Calculation of averages and standard deviations for the ensemble are given by equations (1.1) and (1.2). This pure state $|\psi\rangle$ may be an eigenstate of a particular physical observable, or may be a linear superposition of eigenstates of an arbitrary operator.

Definition 2 *Mixed ensemble.* *The wave function is not known with certainty, and we have several possibilities that we write as $\{|\psi^{(1)}\rangle, |\psi^{(2)}\rangle \dots |\psi^{(i)}\rangle \dots\}$. This collection may be finite or infinite. A fraction of the members of the ensemble, with relative population w_i , is assigned to the ket $|\psi^{(i)}\rangle$. Normalization requires that*

$$\sum_i w_i = 1. \tag{1.3}$$

The weights $\{w_i\}$ are positive (or zero) real numbers. We say that this ensemble represents a macrostate of the system.

CHAPTER 1 — MANUSCRIPT

The states vectors $\left\{ \left| \psi^{(1)} \right\rangle, \left| \psi^{(2)} \right\rangle \dots \left| \psi^{(i)} \right\rangle \dots \right\}$ are normalized but not necessarily orthogonal, and the number of such states may be greater than the dimension of the linear space. The numbers $\{w_i\}$ cannot be interpreted as ordinary probabilities, since the quantum states $\left\{ \left| \psi^{(1)} \right\rangle, \left| \psi^{(2)} \right\rangle \dots \left| \psi^{(i)} \right\rangle \dots \right\}$ are not ‘mutually exclusive’. This has to be understood in the sense that the overlap between two states of the collection does not vanish in general, *i.e.*

$$\left\langle \psi^{(i)} \left| \psi^{(j)} \right\rangle \neq 0 \quad (1.4)$$

in the general case. Now, we have to prescribe the calculation of averages for the mixed ensemble. Let \mathbf{A} be a physical observable. We denote by [...] the average for the ensemble, to distinguish from $\langle \dots \rangle$ used for the pure quantum case.

Definition 3 *Average or Mean Value for the mixed ensemble.*

$$[\mathbf{A}] \equiv \sum_i w_i \left\langle \psi^{(i)} \left| \mathbf{A} \right| \psi^{(i)} \right\rangle . \quad (1.5)$$

In the above definition, we see that the ordinary quantum average $\left\langle \psi^{(i)} \left| \mathbf{A} \right| \psi^{(i)} \right\rangle$ for the state $\psi^{(i)}$ is weighed by its relative population w_i , so the average $[\mathbf{A}]$ has a mixed quantum and statistical nature. We rewrite the definition (1.5) using a general basis of

CHAPTER 1 — MANUSCRIPT

states $\{|n\rangle\}$, which is orthonormal and complete:

$$\begin{aligned} [\mathbf{A}] &= \sum_i w_i \sum_{n,n'} \langle \psi^{(i)} | n \rangle \langle n | \mathbf{A} | n' \rangle \langle n' | \psi^{(i)} \rangle \\ &= \sum_{n,n'} \left(\sum_i w_i \langle n' | \psi^{(i)} \rangle \langle \psi^{(i)} | n \rangle \right) \langle n | \mathbf{A} | n' \rangle . \end{aligned} \quad (1.6)$$

The expression above suggests the following definition:

Definition 4 *State Operator or Density Operator, ρ*

$$\rho \equiv \sum_i w_i |\psi^{(i)}\rangle \langle \psi^{(i)}|. \quad (1.7)$$

Its matrix elements are given by:

$$\langle n' | \rho | n \rangle = \sum_i w_i \langle n' | \psi^{(i)} \rangle \langle \psi^{(i)} | n \rangle ,$$

and the mean value can be written as a trace:

$$[\mathbf{A}] = \sum_{n,n'} \langle n' | \rho | n \rangle \langle n | \mathbf{A} | n' \rangle = \text{Tr}(\rho \mathbf{A}) . \quad (1.8)$$

From the definition, we get some immediate properties:

i) ρ is an Hermitian operator:

$$\rho^\dagger = \sum_i w_i^* \left(|\psi^{(i)}\rangle \langle \psi^{(i)}| \right)^\dagger = \sum_i w_i |\psi^{(i)}\rangle \langle \psi^{(i)}| = \rho ,$$

because the $\{w_i\}$ are real numbers. The eigenvalues of ρ are then real. We will show that they are all positive;

CHAPTER 1 — MANUSCRIPT

ii) the density operator is normalized in the sense:

$$\begin{aligned} \text{Tr } \boldsymbol{\rho} &= \sum_n \sum_i w_i \langle n | \psi^{(i)} \rangle \langle \psi^{(i)} | n \rangle = \sum_i w_i \sum_n \langle \psi^{(i)} | n \rangle \langle n | \psi^{(i)} \rangle = \\ &= \sum_i w_i \langle \psi^{(i)} | \psi^{(i)} \rangle = \sum_i w_i = 1, \end{aligned} \quad (1.9)$$

since the states $\psi^{(i)}$ are normalized. This condition reduces to the fact that the unit operator $\mathbf{1}$ has mean value equal to 1;

iii) consider an observable \mathbf{A} . The operator $\mathbf{A}^\dagger \mathbf{A}$ is said to be positive definite (non negative). That means that its average is positive or zero:

$$[\mathbf{A}^\dagger \mathbf{A}] = \sum_i w_i \langle \psi^{(i)} | \mathbf{A}^\dagger \mathbf{A} | \psi^{(i)} \rangle = \sum_i w_i \langle \varphi^{(i)} | \varphi^{(i)} \rangle,$$

where $|\varphi^{(i)} \rangle = \mathbf{A} |\psi^{(i)} \rangle$. Since the metric is positive, we have $\langle \varphi^{(i)} | \varphi^{(i)} \rangle \geq 0$, resulting $[\mathbf{A}^\dagger \mathbf{A}] \geq 0$. Since \mathbf{A} is an observable, it is Hermitian, $\mathbf{A}^\dagger = \mathbf{A}$, with real eigenvalues. Consider the basis that leaves \mathbf{A} in diagonal form, $A_{mn} = A_n \delta_{mn}$ and calculate $[\mathbf{A}^\dagger \mathbf{A}]$:

$$0 \leq [\mathbf{A}^\dagger \mathbf{A}] = \text{Tr}(\boldsymbol{\rho} \mathbf{A} \mathbf{A}) = \sum_{k,m,n} \rho_{nm} A_{mk} A_{kn} = \sum_n \rho_{nn} A_n^2,$$

and since \mathbf{A} is arbitrary, we obtain that $\rho_{nn} \geq 0$. Result: any diagonal element of $\boldsymbol{\rho}$, for an arbitrary representation, is non negative. In particular, if we choose the representation where the operator $\boldsymbol{\rho}$ is diagonal, $\rho_{mn} = \rho_n \delta_{mn}$, we obtain that the eigenvalues are non negative, $\rho_n \geq 0$, with $\sum_n \rho_n = 1$;

CHAPTER 1 — MANUSCRIPT

iv) using the same representation, we obtain

$$\text{Tr}(\boldsymbol{\rho}^2) = \sum_n \rho_n^2 \leq \left(\sum_n \rho_n \right)^2 = (\text{Tr} \boldsymbol{\rho})^2 = 1 ,$$

that is $\text{Tr}(\boldsymbol{\rho}^2) \leq 1$, and since the trace is invariant, this result is valid in any representation;

v) the pure ensemble can be considered as a limit case of the mixed ensemble, when only one of the weights is different from zero, *i.e.*

$$w_i = \begin{cases} 1, & \text{for } i = j, \\ 0, & \text{for } i \neq j . \end{cases}$$

and $\boldsymbol{\rho} = |\psi^{(j)} \rangle \langle \psi^{(j)}|$. In the following, we will review some properties of the pure ensemble. ■

A 1.2 Pure versus mixed ensembles

For pure states, the density operator has the simple form:

$$\boldsymbol{\rho} = |\psi \rangle \langle \psi| ,$$

from where we get the properties:

a) idempotence of $\boldsymbol{\rho}$,

$$\boldsymbol{\rho}^2 = (|\psi \rangle \langle \psi|)(|\psi \rangle \langle \psi|) = |\psi \rangle \langle \psi| \psi \rangle \langle \psi| = |\psi \rangle \langle \psi| = \boldsymbol{\rho} ,$$

CHAPTER 1 — MANUSCRIPT

since $\langle \psi | \psi \rangle = 1$. This relation can be factorized:

$$\boldsymbol{\rho}(\boldsymbol{\rho} - \mathbf{1}) = 0 . \quad (1.10)$$

b) the equation (1.10) is also satisfied by its eigenvalues:

$$\rho_n(\rho_n - 1) = 0 ,$$

with solutions $\rho_n = 0, 1$. Due to normalization, $\sum_n \rho_n = 1$, it follows that only one of the eigenvalues is 1, and all the other are zero. In the diagonal form, the density matrix is written as

$$\boldsymbol{\rho} \doteq \begin{pmatrix} 0 & 0 & . & . & . & 0 & 0 \\ 0 & \dots & & & & \dots & 0 \\ . & & 0 & & & & . \\ . & & & 1 & & & . \\ . & & & & 0 & & . \\ 0 & \dots & & & & \dots & 0 \\ 0 & 0 & . & . & . & 0 & 0 \end{pmatrix} .$$

c) due to the idempotent property, we get

$$\text{Tr}(\boldsymbol{\rho}^2) = \text{Tr}(\boldsymbol{\rho}) = 1 . \quad (1.11)$$

Relation (1.11) can be considered as a necessary and sufficient condition for an ensemble to be pure.

CHAPTER 1 — MANUSCRIPT

In contrast, for the mixed ensemble we have $\rho^2 \neq \rho$, and $\text{Tr}(\rho^2) < 1$. The decomposition $\rho = \sum_i w_i |\psi^{(i)}\rangle\langle\psi^{(i)}|$ is not unique, as we shall see in the examples.

EXP **Example 1**

We work out an illustrative example for spin $1/2$. The linear space has dimension 2, and the density operator is represented by a (2×2) matrix. Due to hermiticity and normalization, we are left with three real independent parameters to determine ρ . Those parameters can be identified with the three mean values for the average of the spin operator, $[S_x], [S_y], [S_z]$. This is a particular characteristic of spin $1/2$. For convenience, we introduce the Pauli spin operator $\vec{\sigma}$ by

$$\vec{\mathbf{S}} = \frac{\hbar}{2} \vec{\sigma} ,$$

with the standard representation:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \quad (1.12)$$

We note that the three Pauli matrices plus the identity form a basis of the linear space of the complex (2×2) matrices. So, in general we have

$$\rho = \frac{1}{2} m_0 \mathbf{1} + \frac{1}{2} m_x \sigma_x + \frac{1}{2} m_y \sigma_y + \frac{1}{2} m_z \sigma_z = \frac{1}{2} m_0 \mathbf{1} + \frac{1}{2} \vec{\mathbf{m}} \cdot \vec{\sigma} , \quad (1.13)$$

CHAPTER 1 — MANUSCRIPT

with the vector $\vec{\mathbf{m}} = (m_x, m_y, m_z)$ being called *polarization*. Due to hermiticity of ρ and the Pauli matrices, all the coefficients in (1.13) are real. Also, remembering that Pauli matrices are traceless, the normalization condition reads:

$$\text{Tr}(\rho) = 1 = \frac{1}{2} m_0 \text{Tr}(\mathbf{1}) = m_0 ,$$

and we are left with three real independent parameters given by $\vec{\mathbf{m}} = (m_x, m_y, m_z)$, with the density operator written in the general form:

$$\rho = \frac{1}{2} (\mathbf{1} + \vec{\mathbf{m}} \cdot \vec{\sigma}) .$$

Using the algebra associated with the Pauli matrices, one easily finds that

$$\rho^2 = \frac{1}{4} \left[\mathbf{1} + 2\vec{\mathbf{m}} \cdot \vec{\sigma} + (\vec{\mathbf{m}} \cdot \vec{\sigma})^2 \right] = \frac{1}{2} \left[\left(\frac{1+m^2}{2} \right) \mathbf{1} + \vec{\mathbf{m}} \cdot \vec{\sigma} \right] ,$$

with $m^2 = |\vec{\mathbf{m}}|^2 = m_x^2 + m_y^2 + m_z^2$. Taking the trace, we obtain

$$\text{Tr}(\rho^2) = \frac{1+m^2}{2} \leq 1 ,$$

implying that $m^2 \leq 1$, or $0 \leq |\vec{\mathbf{m}}| \leq 1$. We have a pure ensemble if and only if $m^2 = 1$, which means maximum polarization. For the mixed ensemble, $0 \leq m^2 < 1$. The case $m = 0$ is called unpolarized or random ensemble. Using the anticommuting properties of Pauli

CHAPTER 1 — MANUSCRIPT

matrices:

$$\boldsymbol{\sigma}_x \boldsymbol{\sigma}_y = -\boldsymbol{\sigma}_y \boldsymbol{\sigma}_x = i\boldsymbol{\sigma}_z ,$$

$$\boldsymbol{\sigma}_y \boldsymbol{\sigma}_z = -\boldsymbol{\sigma}_z \boldsymbol{\sigma}_y = i\boldsymbol{\sigma}_x ,$$

$$\boldsymbol{\sigma}_z \boldsymbol{\sigma}_x = -\boldsymbol{\sigma}_x \boldsymbol{\sigma}_z = i\boldsymbol{\sigma}_y ,$$

it follows

$$[\mathbf{S}_x] = \text{Tr}(\boldsymbol{\rho} \mathbf{S}_x) = \frac{\hbar}{2} m_x ,$$

$$[S_y] = \text{Tr}(\boldsymbol{\rho} \mathbf{S}_y) = \frac{\hbar}{2} m_y ,$$

$$[S_z] = \text{Tr}(\boldsymbol{\rho} \mathbf{S}_z) = \frac{\hbar}{2} m_z .$$

For the random ensemble, $[S_x] = [S_y] = [S_z] = 0$, that is $m = 0$ and

the density operator is written as:

$$\boldsymbol{\rho}_0 = \frac{1}{2} \mathbf{1} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} . \tag{1.14}$$

We will discover that this case corresponds to the maximal mixing.

Suppose that we use the basis of states that diagonalize \mathbf{S}_z . We call

them $|\hat{\mathbf{z}}; + \rangle$ and $|\hat{\mathbf{z}}; - \rangle$, corresponding to the eigenvalues $\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$ respectively. The density operator for the random case in (1.14)

can be represented in the form

$$\begin{aligned} \boldsymbol{\rho}_0 &= \frac{1}{2} |\hat{\mathbf{z}}; + \rangle \langle \hat{\mathbf{z}}; +| + \frac{1}{2} |\hat{\mathbf{z}}; - \rangle \langle \hat{\mathbf{z}}; -| = \\ &= w_+ |\hat{\mathbf{z}}; + \rangle \langle \hat{\mathbf{z}}; +| + w_- |\hat{\mathbf{z}}; - \rangle \langle \hat{\mathbf{z}}; -| , \end{aligned}$$

CHAPTER 1 — MANUSCRIPT

with $w_+ = w_- = \frac{1}{2}$, *i.e.* ρ_0 can be considered as a mixing of the states $|\hat{\mathbf{z}}; + \rangle$ and $|\hat{\mathbf{z}}; - \rangle$ with equal weights. One important issue for the mixed ensemble, is that the decomposition in terms of pure ensembles is not unique. In the example above, we can consider a different basis of states, let's say the states $|\hat{\mathbf{x}}; + \rangle$ and $|\hat{\mathbf{x}}; - \rangle$ that diagonalize the component \mathbf{S}_x of the spin operator. We have the unitary transformation:

$$\begin{aligned} |\hat{\mathbf{z}}; + \rangle &= \frac{1}{\sqrt{2}} |\hat{\mathbf{x}}; + \rangle + \frac{1}{\sqrt{2}} |\hat{\mathbf{x}}; - \rangle, \\ |\hat{\mathbf{z}}; - \rangle &= \frac{1}{\sqrt{2}} |\hat{\mathbf{x}}; + \rangle - \frac{1}{\sqrt{2}} |\hat{\mathbf{x}}; - \rangle, \end{aligned}$$

that leads to

$$\rho_0 = \frac{1}{2} |\hat{\mathbf{x}}; + \rangle \langle \hat{\mathbf{x}}; +| + \frac{1}{2} |\hat{\mathbf{x}}; - \rangle \langle \hat{\mathbf{x}}; -|,$$

which means that the random ensemble (1.14) may be considered, at the same time, as a mixing of states $|\hat{\mathbf{x}}; + \rangle$ and $|\hat{\mathbf{x}}; - \rangle$ with equal weights. Actually, there is an infinite number of possibilities, saying that the random mixed ensemble can be decomposed equally in terms of black and white, or red and green, or blue and yellow, and so on, at the same time. This fact is a manifestation of the quantum nature of the state, in spite of the maximal mixture. The

CHAPTER 1 — MANUSCRIPT

pure ensemble can be parameterized as:

$$\vec{\mathbf{m}} = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta) ,$$

with $|\tilde{\mathbf{m}}|^2 = 1$, with the angles (θ, φ) giving the direction of the polarization. In matrix form, the density operator reads:

$$\rho = \begin{pmatrix} \frac{1+\cos \theta}{2} & \frac{e^{-i\varphi} \sin \theta}{2} \\ \frac{e^{i\varphi} \sin \theta}{2} & \frac{1-\cos \theta}{2} \end{pmatrix} .$$

For many other examples, see the exercise section. ■

A

1.3 Coupled systems and non-separability

Two quantum systems that interacted in the past, remain correlated forever and cannot be represented separately by ket states. In plain words, a subsystem can not in general, be described by a wave function. This feature is called the *Non-separability Property*. This is a limitation of the wave function formalism, since in Statistical Mechanics, one usually wants to refer to a part of the whole system, independently of the rest. Important cases are those when the system is coupled to *reservoirs*, and we want to eliminate the latter degrees of freedom from the physical description. Those are instances where the density operator furnishes a superior view of the state of the system, since the wave function representation has been ruled out. To fully understand the non-

CHAPTER 1 — MANUSCRIPT

separable case, we will firstly discuss the case of non-interacting systems, that is systems that are separable.

A diluted system, as for example an *ideal gas*, can be well approximated as a system of many non-interacting particles (molecules). Within this ideal limit, particles are not correlated, and the Statistical Mechanics can be reduced to a *one-particle* description. In a general way, let us consider two systems which span two different Hilbert spaces of ket states, \mathcal{R} and \mathcal{S} , with basis $\{|N \rangle\}$ and $\{|n \rangle\}$, respectively. The combined space of the two systems, $\mathcal{R} \times \mathcal{S}$, is spanned by the direct product states, which in the Dirac's notation, are written as:

$$|Nn \rangle \equiv |N \rangle |n \rangle .$$

Direct products of operators are represented by direct products of the corresponding matrices.

Example. Direct product of two matrices A and B , where:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} .$$

One defines $A \times B$ as the matrix:

$$A \times B \equiv \begin{pmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{pmatrix} .$$

CHAPTER 1 — MANUSCRIPT

The direct product results in a (6×6) matrix, since the terms $a_{ij}B$ are understood as (3×3) submatrices.

The following properties are easily demonstrated:

1. Matrix elements are understood as:

$$\langle Nn|\mathbf{A} \times \mathbf{B}|N'n' \rangle = \langle N|\mathbf{A}|N' \rangle \langle n|\mathbf{B}|n' \rangle .$$

In particular, scalar products are obtained as:

$$\langle Nn|N'n' \rangle = \langle N|N' \rangle \langle n|n' \rangle = \delta_{NN'} \delta_{nn'} .$$

2. From the point above, we get:

$$\text{Tr}(\mathbf{A} \times \mathbf{B}) = \text{Tr}(\mathbf{A}) \text{Tr}(\mathbf{B}) .$$

3. Direct products of column (row) vectors are obtained using the same rule defined for matrices. Take for example the case below:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ 0 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ 0 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} .$$

CHAPTER 1 — MANUSCRIPT

A general state of the system is written as:

$$|\psi\rangle = \sum_{N,n} \mathfrak{C}(N,n) |N\rangle |n\rangle ,$$

but since the systems are uncorrelated, one must have $\mathfrak{C}(N,n) = C(N)c(n)$ and the state is separable:

$$|\psi\rangle = \sum_{N,n} \mathfrak{C}(N,n) |N\rangle |n\rangle = \left(\sum_N C(N) |N\rangle \right) \left(\sum_n c(n) |n\rangle \right) = |\psi_R\rangle |\psi_S\rangle .$$

The same is true for the density matrix, which in general is written as

$$\rho = \sum_{N,M,n,m} |Nn\rangle \langle Nn| \rho |Mm\rangle \langle Mm| ,$$

but for uncorrelated systems, one should have

$$\langle Nn | \rho | Mm \rangle = A_{NM} B_{nm} ,$$

factorizing the density operator as:

$$\rho = \left(\sum_N A_{NM} |N\rangle \langle M| \right) \left(\sum_n B_{nm} |n\rangle \langle m| \right) = \rho_{\mathcal{R}} \times \rho_{\mathcal{S}} , \quad (1.15)$$

in the form of a direct product. Separability for the density operator has a broader sense in Quantum Mechanics, but we shall not pursue this discussion here [4]. A situation as the one depicted in (1.15), is called *simple separability*. In the case of an ideal gas, in the absence of interactions, all the particles are uncorrelated. If we use the particle coordinates as labels, separability leads to:

$$\rho(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \rho_1(\mathbf{x}_1) \times \rho_2(\mathbf{x}_2) \times \dots \times \rho_N(\mathbf{x}_N) ,$$

CHAPTER 1 — MANUSCRIPT

where N is the total number of particles and $\rho_i(\mathbf{x}_i)$ is the density operator for one-particle states. If the particles are identical,

$$\rho_j = \rho_1,$$

for all $j = 2, 3, \dots, N$, and the factorization is written as

$$\rho(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \bigotimes_{i=1}^N \rho_1(\mathbf{x}_i) .$$

In particular, if ρ_1 is normalized, we obtain:

$$\text{Tr} [\rho(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)] = \prod_{i=1}^N \text{Tr} [\rho_1(\mathbf{x}_i)] = 1 .$$

In this case, the calculation of the density operator is reduced to the calculation of the one-particle operator $\rho_1(\mathbf{x})$.

In the interacting case, correlations among the particles appear, and the states are not any more separable. Assume that at the initial time ($t_0 = 0$), the two subsystem are not interacting, and we prepare the initial state as separable:

$$|\psi, 0 \rangle = |N \rangle |n \rangle .$$

After that, the interaction is turned on during a finite interval of time, and finally turned off again. The interaction made the systems to be correlated. The asymptotic state ($t \rightarrow \infty$) is of the type:

$$|\psi, \infty \rangle = \sum_{M,m} a_{\infty}(M, m; N, n) |M \rangle |m \rangle , \quad (1.16)$$

CHAPTER 1 — MANUSCRIPT

where the coefficient $a_\infty(M, m; N, n)$ is the probability amplitude for the transition $|N \rangle |n \rangle \rightarrow |M \rangle |m \rangle$. Since the systems are correlated, we have in general that

$$a_\infty(M, m; N, n) \neq C(M, N)c(m, n) ,$$

for all pairs (M, m) and (N, n) , and the state (1.16) is not separable. In this case, we cannot assign a wave function to a subsystem (either R or S). This result is called *nonseparability principle*.

A

1.4 Density matrix of a subsystem

We have seen in the previous subsection that the language of ket states does not allow us, in general, to describe an isolated system from the remainder of the universe. But this is possible when one describes the state of the system through the density operator. Consider two interacting quantal systems, whose states span the spaces R and S . We want to pay attention to subsystem S , leaving R undetected (R may be a reservoir, and we want to eliminate its degrees of freedom). The states of the total system span $R \times S$, but in general, physical states are not separable in the presence of interactions. Assume basis $\{|N \rangle\}$ and $\{|n \rangle\}$ for R and S , respectively, as in the subsection above. We want to calculate an average of an observable that refers only to S , which is written in the form

$$\Omega = \mathbf{I}_R \times \Omega_S ,$$

CHAPTER 1 — MANUSCRIPT

where I_R is the identity in R :

$$[\mathbf{\Omega}] = \text{Tr} [\boldsymbol{\rho} (\mathbf{I}_R \times \mathbf{\Omega}_S)] = \sum_{N,M,n,m} \langle Nn | \boldsymbol{\rho} | Mm \rangle \langle Mm | (\mathbf{I}_R \times \mathbf{\Omega}_S) | Nn \rangle . \quad (1.17)$$

Note that $\langle Mm | (\mathbf{I}_R \times \mathbf{\Omega}_S) | Nn \rangle = \delta_{MN} \langle m | \mathbf{\Omega}_S | n \rangle$, and substituting in (1.17) yields:

$$[\mathbf{\Omega}] = \sum_{n,m} \langle m | \mathbf{\Omega}_S | n \rangle \left(\sum_N \langle Nn | \boldsymbol{\rho} | Nm \rangle \right) .$$

The quantity $\sum_N \langle Nn | \boldsymbol{\rho} | Nm \rangle$ is the partial trace of $\boldsymbol{\rho}$ in the R space. We then make the following definition:

Definition 5 *Reduced density operator, $\bar{\boldsymbol{\rho}}_S$, relative to S .*

Its matrix elements in S are given by:

$$\langle n | \bar{\boldsymbol{\rho}}_S | m \rangle \equiv \sum_N \langle Nn | \boldsymbol{\rho} | Nm \rangle . \quad (1.18)$$

We rewrite the above definition (1.18) in a formal fashion as:

$$\bar{\boldsymbol{\rho}}_S = \text{Tr}_R (\boldsymbol{\rho}) ,$$

meaning that $\bar{\boldsymbol{\rho}}_S$ is obtained from $\boldsymbol{\rho}$ by taking the partial trace in R . The average value (1.17) is now referred to space S only:

$$[\mathbf{\Omega}] = \text{Tr} [\boldsymbol{\rho} (\mathbf{I}_R \times \mathbf{\Omega}_S)] = \text{Tr}_S [\bar{\boldsymbol{\rho}}_S \mathbf{\Omega}_S] .$$

By taking the partial trace, we lose the detailed information relative to subsystem R .

It remains to prove that $\bar{\boldsymbol{\rho}}_S$ is a *bona fide* density operator. This is achieved, if we assume that $\boldsymbol{\rho}$ is a density operator for the whole universe $R \times S$:

CHAPTER 1 — MANUSCRIPT

i) hermiticity,

$$\langle n|\bar{\rho}_S|m\rangle \equiv \sum_N \langle Nn|\rho|Nm\rangle = \sum_N \langle Nm|\rho|Nn\rangle^* = \langle m|\bar{\rho}_S|n\rangle^*,$$

because ρ is Hermitian;

ii) normalization,

$$\text{Tr}_S(\bar{\rho}_S) = \sum_n \langle n|\bar{\rho}_S|n\rangle \equiv \sum_n \sum_N \langle Nn|\rho|Nn\rangle = \text{Tr}_S[\text{Tr}_R(\rho)] = \text{Tr}(\rho) = 1;$$

iii) positivity,

$$\langle n|\bar{\rho}_S|n\rangle \equiv \sum_N \langle Nn|\rho|Nn\rangle \geq 0,$$

since it is a sum of positive terms.

To measure properties of the S subsystem we do not need the complete density operator ρ , but only the reduced operator $\bar{\rho}_S$ relative to S . Detailed information of the other subsystem is lost, but some ‘average properties’ of R are still contained in $\bar{\rho}_S$.

EXP **Example 2**

Consider two interacting particles of spin 1/2, which are coupled in a singlet state of the total spin:

$$|\psi_0\rangle = \frac{1}{\sqrt{2}} |+\rangle_R |-\rangle_S - \frac{1}{\sqrt{2}} |-\rangle_R |+\rangle_S, \quad (1.19a)$$

where we have used the labels R and S for the particles. The above state is said to be *entangled*, and clearly there is no ket state to

CHAPTER 1 — MANUSCRIPT

represent either subsystem R or S . The state (1.19a) is a pure state, with density operator given by:

$$\begin{aligned} \rho &= |\psi_0\rangle\langle\psi_0| = \\ &= \frac{1}{2}\{|+\rangle_R|-\rangle_S \langle -|_R \langle +| + |-\rangle_R|+\rangle_S \langle +|_R \langle -| - \\ &\quad - |+\rangle_R|-\rangle_S \langle +|_R \langle -| - |-\rangle_R|+\rangle_S \langle -|_R \langle +|\} . \end{aligned}$$

We now take the partial trace relative to R :

$$\begin{aligned} \bar{\rho}_S &= \text{Tr}_R(\rho) = {}_R\langle +|\rho|+\rangle_R + {}_R\langle -|\rho|-\rangle_R = \\ &= \frac{1}{2}|-\rangle_S \langle -| + \frac{1}{2}|+\rangle_S \langle +| \doteq \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_S . \end{aligned} \quad (1.20)$$

Note that $\bar{\rho}_S$ represents a mixed ensemble (maximum mixture), while the original ρ was a pure state: by eliminating the degrees of freedom of R , some information is lost in a way that cannot be recovered later. Just to stress the irreversible character of the process, we note that partially tracing the density matrix of the pure ensemble of the triplet state

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}|+\rangle_R|-\rangle_S + \frac{1}{\sqrt{2}}|-\rangle_R|+\rangle_S$$

leads to the same reduced density matrix (1.20). ■

A

1.5 Representing the Density Operator: Density Matrix

We discuss in first place the case of discrete spectrum, with a complete and orthonormal basis $\{|n\rangle\}$. In many instances, $\{|n\rangle\}$ is the basis that makes diagonal a complete set of observables. We expand the density operator in terms of this basis :

$$\begin{aligned} \rho &= \sum_i w_i |\psi^{(i)}\rangle\langle\psi^{(i)}| = \sum_{n,m} \sum_i w_i |n\rangle\langle n|\psi^{(i)}\rangle\langle\psi^{(i)}|m\rangle\langle m| = \\ &= \sum_{n,m} |n\rangle\langle n|\rho|m\rangle\langle m| = \sum_{n,m} |n\rangle\langle m| \left(\sum_i w_i \langle n|\psi^{(i)}\rangle\langle\psi^{(i)}|m\rangle \right) . \end{aligned}$$

Defining the linear coefficients as $a_n^{(i)} \equiv \langle n|\psi^{(i)}\rangle$, *i.e.* $|\psi^{(i)}\rangle = \sum_n a_n^{(i)}|n\rangle$, the matrix elements of ρ are written as:

$$\langle n|\rho|m\rangle = \sum_i w_i a_n^{(i)} a_m^{(i)*} \equiv \overline{a_n a_m^*} ,$$

where the bar means average over the mixed ensemble and a^* is the complex conjugate of a . We remember that $a_n^{(i)} \equiv \langle n|\psi^{(i)}\rangle$ is the probability amplitude that the state $|n\rangle$ is contained in $|\psi^{(i)}\rangle$. The corresponding probability is $P_n^{(i)} = |a_n^{(i)}|^2 = |\langle n|\psi^{(i)}\rangle|^2$, which appears in the diagonal matrix elements of ρ :

$$\langle n|\rho|n\rangle = \sum_i w_i |a_n^{(i)}|^2 = \sum_i w_i P_n^{(i)} \geq 0 ,$$

which in turn, can be interpreted as a probability, since:

$$\sum_n \langle n|\rho|n\rangle = \sum_i w_i \sum_n P_n^{(i)} = \sum_i w_i = 1 .$$

CHAPTER 1 — MANUSCRIPT

In other words, the diagonal matrix element $\langle n|\rho|n\rangle$ yields the probability that the state $|n\rangle$ is occupied in the ensemble represented by ρ . We can also refer the average of a physical quantity \mathbf{A} to the same basis:

$$[\mathbf{A}] = \text{Tr}(\rho\mathbf{A}) = \sum_{n,m} \langle n|\rho|m\rangle \langle m|\mathbf{A}|n\rangle = \sum_{n,m} \sum_i w_i a_n^{(i)} a_m^{(i)*} A_{mn} = \sum_{n,m} \overline{a_n a_m^*} A_{mn} .$$

If an operator \mathbf{B} is diagonal in the given basis, *i.e.* $\langle m|\mathbf{B}|n\rangle = B_n \delta_{mn}$, its average is given by:

$$[\mathbf{B}] = \sum_i w_i \sum_n P_n^{(i)} B_n = \sum_n B_n \overline{|a_n|^2} .$$

For the continuous spectrum, we discuss the important cases of the coordinate and momentum representations. The matrix elements of ρ now become two-point functions. For the coordinate representation $\{|\vec{\mathbf{x}}\rangle\}$, the linear coefficients $\langle \vec{\mathbf{x}}|\psi^{(i)}\rangle$ are called ‘wave functions’ and usually, one writes

$$\psi^{(i)}(\vec{\mathbf{x}}) = \langle \vec{\mathbf{x}}|\psi^{(i)}\rangle ,$$

meaning that they vary continuously with $\vec{\mathbf{x}}$. Now, we represent the density operator:

$$\begin{aligned} \langle \vec{\mathbf{x}}|\rho|\vec{\mathbf{x}}'\rangle &= \rho(\vec{\mathbf{x}}, \vec{\mathbf{x}}') = \sum_i w_i \langle \vec{\mathbf{x}}|\psi^{(i)}\rangle \langle \psi^{(i)}|\vec{\mathbf{x}}'\rangle = \\ &= \sum_i w_i \psi^{(i)}(\vec{\mathbf{x}}) \psi^{(i)*}(\vec{\mathbf{x}}') = \overline{\psi(\vec{\mathbf{x}}) \psi^*(\vec{\mathbf{x}}')} . \end{aligned} \quad (1.21)$$

Diagonal elements are given by:

$$\langle \vec{\mathbf{x}}|\rho|\vec{\mathbf{x}}\rangle = \rho(\vec{\mathbf{x}}, \vec{\mathbf{x}}) = \sum_i w_i \left| \psi^{(i)}(\vec{\mathbf{x}}) \right|^2 = \overline{|\psi(\vec{\mathbf{x}})|^2} \geq 0 . \quad (1.22)$$

CHAPTER 1 — MANUSCRIPT

In Quantum Mechanics, $\left| \psi^{(i)}(\vec{\mathbf{x}}) \right|^2$ is interpreted as a probability density, since normalization requires

$$\int d\vec{\mathbf{x}} \left| \psi^{(i)}(\vec{\mathbf{x}}) \right|^2 = 1 ,$$

where the integral is taken over all space. It follows that diagonal elements of ρ , given by (1.22), are also probability densities, with

$$\text{Tr}(\rho) = \int d\vec{\mathbf{x}} \langle \vec{\mathbf{x}} | \rho | \vec{\mathbf{x}} \rangle = \sum_i w_i = 1 .$$

Average of an observable \mathbf{A} is obtained integrating the two-point function $\rho(\vec{\mathbf{x}}, \vec{\mathbf{x}}')$ of (1.21) with the matrix elements of \mathbf{A} :

$$\begin{aligned} [\mathbf{A}] &= \text{Tr}(\rho \mathbf{A}) = \int d\vec{\mathbf{x}} \int d\vec{\mathbf{x}}' \langle \vec{\mathbf{x}} | \rho | \vec{\mathbf{x}}' \rangle \langle \vec{\mathbf{x}}' | \mathbf{A} | \vec{\mathbf{x}} \rangle = \\ &= \int d\vec{\mathbf{x}} \int d\vec{\mathbf{x}}' \rho(\vec{\mathbf{x}}, \vec{\mathbf{x}}') A(\vec{\mathbf{x}}', \vec{\mathbf{x}}) . \end{aligned} \quad (1.23)$$

If the observable is local in the coordinate representation, $A(\vec{\mathbf{x}}', \vec{\mathbf{x}}) = A(\vec{\mathbf{x}}) \delta(\vec{\mathbf{x}}' - \vec{\mathbf{x}})$, where $\delta(\vec{\mathbf{x}}' - \vec{\mathbf{x}})$ is the Dirac delta function, the double integral (1.23) is reduced to the single integration below:

$$[\mathbf{A}] = \int d\vec{\mathbf{x}} \rho(\vec{\mathbf{x}}, \vec{\mathbf{x}}) A(\vec{\mathbf{x}}) ,$$

i.e. the function $A(\vec{\mathbf{x}})$ is integrated with the probability density $\rho(\vec{\mathbf{x}}, \vec{\mathbf{x}})$. Note that for a pure ensemble, $\rho(\vec{\mathbf{x}}, \vec{\mathbf{x}})$ is simply given by $\psi^*(\vec{\mathbf{x}}) \psi(\vec{\mathbf{x}}) = |\psi(\vec{\mathbf{x}})|^2$, the probability density associated with the wave function $\psi(\vec{\mathbf{x}})$.

CHAPTER 1 — MANUSCRIPT

We pass to the momentum representation replacing the wave function in real space by the wave function in momentum space:

$$\phi^{(i)}(\vec{\mathbf{p}}) = \langle \vec{\mathbf{p}} | \psi^{(i)} \rangle .$$

All formulae can be translated to the new basis $\{|\vec{\mathbf{p}}\rangle\}$. For instance, the probability density associated with the diagonal matrix elements of ρ is obtained as:

$$\langle \vec{\mathbf{p}} | \rho | \vec{\mathbf{p}} \rangle = \rho(\vec{\mathbf{p}}, \vec{\mathbf{p}}) = \sum_i w_i \left| \phi^{(i)}(\vec{\mathbf{p}}) \right|^2 = \overline{|\phi(\vec{\mathbf{p}})|^2} \geq 0 .$$

Due to the uncertainty principle, we are bound to use one basis only, but we can go from one to the other, say from $|\vec{\mathbf{x}}\rangle$ to $|\vec{\mathbf{p}}\rangle$, with the transformation matrix [1], [2] :

$$\langle \vec{\mathbf{p}} | \vec{\mathbf{x}} \rangle = \frac{1}{(2\pi\hbar)^{3/2}} \exp\left(-\frac{i\vec{\mathbf{p}} \cdot \vec{\mathbf{x}}}{\hbar}\right) . \quad (1.24)$$

The classical density function that yields the density of points in phase space, is a function of generalized coordinates and momenta, $\rho(q, p)$. To make connection with the classical case, one has to look for a mixed representation of the operator ρ . At first sight, this may seem forbidden by the uncertainty principle, but one can follow a procedure due to Wigner [5] to generate such a function.

Take for instance $\rho(\vec{\mathbf{p}}, \vec{\mathbf{p}})$ and use the transformation (1.24) to pass to the coordinate representation:

$$\rho(\vec{\mathbf{p}}, \vec{\mathbf{p}}) = \int d\vec{\mathbf{x}} \int d\vec{\mathbf{x}}' \frac{1}{(2\pi\hbar)^3} \exp\frac{i\vec{\mathbf{p}} \cdot (\vec{\mathbf{x}}' - \vec{\mathbf{x}})}{\hbar} \rho(\vec{\mathbf{x}}, \vec{\mathbf{x}}') ,$$

CHAPTER 1 — MANUSCRIPT

then make the change of variables with unit Jacobian:

$$\begin{aligned}\vec{\mathbf{r}} &= \vec{\mathbf{x}}' - \vec{\mathbf{x}} , \\ \vec{\mathbf{R}} &= \frac{1}{2} (\vec{\mathbf{x}} + \vec{\mathbf{x}}') .\end{aligned}$$

We obtain

$$\rho(\vec{\mathbf{p}}, \vec{\mathbf{p}}) = \int d\vec{\mathbf{R}} \int d\vec{\mathbf{r}} \frac{1}{(2\pi\hbar)^3} \exp\left(\frac{i\vec{\mathbf{p}} \cdot \vec{\mathbf{r}}}{\hbar}\right) \rho\left(\vec{\mathbf{R}} - \frac{\vec{\mathbf{r}}}{2}, \vec{\mathbf{R}} + \frac{\vec{\mathbf{r}}}{2}\right) . \quad (1.25)$$

The integrand in relation to $\vec{\mathbf{R}}$ defines a function which depends on $\vec{\mathbf{R}}$ and $\vec{\mathbf{p}}$, which is of the mixed form and consistent with the uncertainty principle.

Definition 6 Wigner function, ρ_W

$$\rho_W(\vec{\mathbf{x}}, \vec{\mathbf{p}}) \equiv \frac{1}{(2\pi\hbar)^3} \int d\vec{\mathbf{r}} \exp\left(\frac{i\vec{\mathbf{p}} \cdot \vec{\mathbf{r}}}{\hbar}\right) \rho\left(\vec{\mathbf{x}} - \frac{\vec{\mathbf{r}}}{2}, \vec{\mathbf{x}} + \frac{\vec{\mathbf{r}}}{2}\right) \quad (1.26)$$

In our deduction in (1.25), we have proved an important property of the Wigner function, which reads:

$$\rho(\vec{\mathbf{p}}, \vec{\mathbf{p}}) = \int d\vec{\mathbf{x}} \rho_W(\vec{\mathbf{x}}, \vec{\mathbf{p}}) , \quad (1.27)$$

that is, the probability density $\rho(\vec{\mathbf{p}}, \vec{\mathbf{p}})$ is obtained from the Wigner function $\rho_W(\vec{\mathbf{x}}, \vec{\mathbf{p}})$ integrating over the space variable $\vec{\mathbf{x}}$. The complementary relation is also true, and it is left to the reader to prove its validity:

$$\rho(\vec{\mathbf{x}}, \vec{\mathbf{x}}) = \int d\vec{\mathbf{p}} \rho_W(\vec{\mathbf{x}}, \vec{\mathbf{p}}) . \quad (1.28)$$

CHAPTER 1 — MANUSCRIPT

Relations (1.27) and (1.28) are desirable properties for a function to be candidate to represent the classical density. But definition (1.26) is fully quantum mechanical, and one has to take the limit in some non trivial way to get the classical case. In general, one can show that the property

$$\rho_W(\vec{\mathbf{x}}, \vec{\mathbf{p}}) \geq 0$$

is not always satisfied in the whole phase space $(\vec{\mathbf{x}}, \vec{\mathbf{p}})$. Regions where $\rho_W(\vec{\mathbf{x}}, \vec{\mathbf{p}}) < 0$ are said to contain coherent quantum effects, the size of those regions shrinking with $\hbar \rightarrow 0$. Formally, it was shown by Wigner [5] that ρ_W satisfies the Liouville equation, when $\hbar \rightarrow 0$ (see next subsection). For a system consisting of N particles, definition (1.26) can be generalized to the phase space Γ of a system of particles:

$$\begin{aligned} \rho_W^{(N)}(\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \dots, \tilde{\mathbf{x}}_N; \tilde{\mathbf{p}}_1, \tilde{\mathbf{p}}_2, \dots, \tilde{\mathbf{p}}_N) &\equiv \frac{1}{(2\pi\hbar)^3} \int \int \dots \int \mathbf{d}^3\mathbf{r}_1 \mathbf{d}^3\mathbf{r}_2 \dots \mathbf{d}^3\mathbf{r}_N \times \\ &\times \exp\left(i \frac{\tilde{\mathbf{p}}_1 \cdot \tilde{\mathbf{r}}_1 + \dots + \tilde{\mathbf{p}}_N \cdot \tilde{\mathbf{r}}_N}{\hbar}\right) \langle \tilde{\mathbf{x}}_1 - \frac{\tilde{\mathbf{r}}_1}{2}, \dots, \tilde{\mathbf{x}}_N - \frac{\tilde{\mathbf{r}}_N}{2} | \boldsymbol{\rho} | \tilde{\mathbf{x}}_1 + \frac{\tilde{\mathbf{r}}_1}{2}, \dots, \tilde{\mathbf{x}}_N + \frac{\tilde{\mathbf{r}}_N}{2} \rangle . \end{aligned} \quad (1.29)$$

B 1.5.1 Digression over the Gibbsian ensemble

We discuss here some key concepts concerning the classical density function $\rho(q, p)$, which was the goal of Wigner's approach. In classical Statistical Mechanics, a microstate of a system of N particles is represented by a point in 'phase space' Γ . This is

CHAPTER 1 — MANUSCRIPT

a $6N$ dimensional space spanned by the generalized coordinates and conjugate momenta $\{q_i, p_i\}$, that characterize the system. If we use cartesian coordinates,

$$\{q_i\}_{3N} = (x_1, y_1, z_1, \dots, x_N, y_N, z_N) .$$

From the macroscopic point of view, we operate with a reduced set of quantities that are compatible with a great number of microstates, as for example the case of a gas occupying a given volume, at standard conditions of temperature and pressure. In Statistical Mechanics, we are not interested in the detailed motion of a system of many particles, but we just want to calculate some average properties (thermodynamic quantities) that we compute using an *ensemble* of ideal replicas of the same system. The ensemble is then represented by a swarm of points in Γ space, being the points interpreted as different microstates corresponding to different initial conditions of the system, all satisfying the same macroscopic constraints. This idea was introduced by Gibbs at the foundations of Statistical Mechanics, with the relevant quantity to characterize the ensemble being the density of points in Γ space. Let $\rho(q_i, p_i, t)$ be such a distribution, with the notation meaning that ρ depends on all generalized coordinates and momenta, and may also depend explicitly on time. In other words,

$$\rho(q_i, p_i, t) dq^{3N} dp^{3N}$$

is the number of representative system points (microstates) contained at time t in the infinitesimal volume $d\Omega = dq^{3N} dp^{3N}$, with $d\Omega$ centered about the point $\{q_i, p_i\}$

CHAPTER 1 — MANUSCRIPT

in phase space. Those representative points of the ensemble evolve in time, tracing a trajectory in Γ space which is closed (periodic motion) or never autointersects itself. Also, trajectories of different representative points never intersect, since they represent motions with different initial conditions (if two trajectories intersect at a given point, that common point may be chosen as a new initial condition, and the two trajectories should coincide at all times). In Classical Mechanics, one can show that time evolution is a canonical transformation [6], and volume of phase space is a canonical invariant (one of Poincaré's integral invariants). The number of representative points of the ensemble contained in any infinitesimal volume $d\Omega$ is also constant in time. The form of the infinitesimal element changes, but its volume is constant. No representative point can intersect the boundary of $d\Omega$ at any time (same argument as given above). We then can enunciate this result as a theorem:

Theorem 7 (Liouville) *The density $\rho(q, p, t)$ is constant in time, or*

$$\frac{d\rho}{dt} = 0 . \tag{1.30}$$

We can rewrite the theorem (1.30) in a different form:

$$0 = \frac{d\rho}{dt} = \frac{\partial\rho}{\partial t} + \sum_i \left(\dot{q}_i \frac{\partial\rho}{\partial q_i} + \dot{p}_i \frac{\partial\rho}{\partial p_i} \right) , \tag{1.31}$$

showing that the explicit dependence on time is cancelled by the implicit dependence through coordinates and momenta. For a Hamiltonian system, Hamilton equations of

CHAPTER 1 — MANUSCRIPT

motion are satisfied, with Hamiltonian H :

$$\dot{q}_i = \frac{\partial H}{\partial p_i} , \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} , \quad (1.32)$$

which we substitute in (1.31), yielding:

$$0 = \frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \sum_i \left(\frac{\partial \rho}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial \rho}{\partial p_i} \frac{\partial H}{\partial q_i} \right) ,$$

which is written in turn, in term of a Poisson bracket [6] as:

$$0 = \frac{\partial \rho}{\partial t} + \{\rho, H\} , \quad (1.33)$$

with the Poisson bracket defined as $\{A, B\} \equiv \sum_i \left(\frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial B}{\partial q_i} \right)$. Relation (1.33)

is another way to state the Liouville theorem. It can be interpreted geometrically

[**huang**]: the motion of representative points in Γ space resembles closely the motion of

an incompressible fluid. In fact, relation (1.33) has the form of a continuity equation,

if one defines a current density for the flux of points as:

$$\vec{\mathbf{j}} \equiv \rho \vec{\mathbf{v}} ,$$

with the velocity vector written as

$$\vec{\mathbf{v}} = (\dot{q}_1, \dot{q}_2, \dots, \dot{q}_{3N}; \dot{p}_1, \dot{p}_2, \dots, \dot{p}_{3N}) .$$

Due to Hamilton equations of motion (1.32), we get

$$0 = \frac{\partial \rho}{\partial t} + \{\rho, H\} = \frac{\partial \rho}{\partial t} + \nabla \cdot \vec{\mathbf{j}} ,$$

CHAPTER 1 — MANUSCRIPT

with the ‘nabla’ operator defined in Γ space as:

$$\nabla \equiv \left(\frac{\partial}{\partial q_1}, \frac{\partial}{\partial q_2}, \dots, \frac{\partial}{\partial q_{3N}}; \frac{\partial}{\partial p_1}, \frac{\partial}{\partial p_2}, \dots, \frac{\partial}{\partial p_{3N}} \right) .$$

Then, local variations of ρ are caused by the flux of the density current $\vec{\mathbf{j}}$, in any neighborhood of representative points of the ensemble. Let $A(q, p)$ be a dynamical quantity of the system of particles. At the macroscopic level, the value of A that we observed is supposed to be the average over the ensemble, calculated as

$$[A]_C(t) = \frac{\int dq^{3N} dp^{3N} \rho(q, p, t) A(q, p)}{\int dq^{3N} dp^{3N} \rho(q, p, t)} ,$$

with the distribution $\rho(q, p, t)$ satisfying Liouville theorem, and the symbol $[...]_C$ standing for the classical average. In principle, the time dependence of $[A]_C(t)$ should approach its equilibrium value at the stationary situation:

$$\frac{\partial \rho}{\partial t} = \{\rho, H\} = 0 .$$

A stationary density distribution $\rho(q, p)$ will only depend on time-independent integrals of the equations of motion. The simplest assumption is to postulate that ρ is a function of the total energy $H = E$, which is a conserved quantity. The equal a priori probability distribution:

$$\rho = \rho(E) = \begin{cases} \text{constant, if } E - \frac{1}{2}\Delta < H < E + \frac{1}{2}\Delta, \\ 0, & \text{otherwise} \end{cases}$$

CHAPTER 1 — MANUSCRIPT

is called *microcanonical ensemble*, and represents an isolated system. The quantity Δ is chosen, such that $\Delta \ll E$, and is introduced for convenience in the counting of states. In the thermodynamic limit, macroscopic quantities are independent of Δ .

The question of how the system approaches such equilibrium state, is at the heart of Statistical Mechanics, being one of the central problem in Physics since the time of Boltzmann. We will discuss those issues in the next chapter. ■

Bibliography

- [1] J. J. Sakurai, *Modern Quantum Mechanics*, revised ed. (Addison-Wesley, 1994).
- [2] L. E. Ballentine, *Quantum Mechanics, A Modern Development* (World Scientific, 1998).
- [3] U. Fano, *Description of States in Quantum Mechanics by Density Matrix and Operator Techniques*, Rev. Mod. Phys. **29**, 74 (1957).
- [4] Michael A. Nielsen and Isaac L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).
- [5] E. Wigner, *On the Quantum Correction For Thermodynamic Equilibrium*, Phys. Rev. **40**, 749 (1932).
- [6] H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics*, 3rd. ed. (Addison-Wesley, 2002).