

## 2.1: Statistical ensemble and probability

As has been already discussed in Sec. 1.1, statistical physics deals with situations when either unknown initial conditions, or system's complexity, or the laws of its motion (as in the case of quantum mechanics) do not allow a definite prediction of measurement results. The main formalism for the analysis of such systems is the probability theory, so let me start with a very brief review of its basic concepts, using an informal “physical” language – less rigorous but (hopefully) more transparent than standard mathematical treatments,<sup>1</sup> and quite sufficient for our purposes.

Consider  $N \gg 1$  independent similar experiments carried out with *apparently* similar systems (i.e. systems with identical *macroscopic* parameters such as volume, pressure, etc.), but still giving, by any of the reasons listed above, different results of measurements. Such a collection of experiments, together with a fixed method of result processing, is a good example of a *statistical ensemble*. Let us start from the case when the experiments may have  $M$  different *discrete* outcomes, and the number of experiments giving the corresponding different results is  $N_1, N_2, \dots, N_M$ , so that

$$\sum_{m=1}^M N_m = N. \quad (2.1.1)$$

The *probability* of each outcome, for the given statistical ensemble, is then defined as

**Probability:**

$$W_m \equiv \lim_{N \rightarrow \infty} \frac{N_m}{N}. \quad (2.1.2)$$

Though this definition is so close to our everyday experience that it is almost self-evident, a few remarks may still be relevant.

First, the probabilities  $W_m$  depend on the exact statistical ensemble they are defined for, notably including the method of result processing. As the simplest example, consider throwing the standard cubic-shaped dice many times. For the ensemble of all thrown and counted dice, the probability of each outcome (say, “1”) is 1/6. However, nothing prevents us from defining another statistical ensemble of dice-throwing experiments in which all outcomes “1” are discounted. Evidently, the probability of finding outcomes “1” in this modified (but legitimate) ensemble is 0, while for all other five outcomes (“2” to “6”), it is 1/5 rather than 1/6.

$$\langle N_m \rangle \equiv W_m N, \quad (2.1.3)$$

with the relative deviations decreasing as  $\sim 1/\langle N_m \rangle^{1/2}$ , i.e. as  $1/N^{1/2}$ .

Now let me list those properties of probabilities that we will immediately need. First, dividing both sides of Equation (2.1.1) by  $N$  and following the limit  $N \rightarrow \infty$ , we get the well-known *normalization condition*

$$\sum_{m=1}^M W_m = 1; \quad (2.1.4)$$

just remember that it is true only if each experiment definitely yields one of the outcomes  $N_1, N_2, \dots, N_M$ .

Second, if we have an additive function of the results,

$$f = \frac{1}{N} \sum_{m=1}^M N_m f_m, \quad (2.1.5)$$

where  $f_m$  are some definite (deterministic) coefficients, the *statistical average* (also called the *expectation value*) of the function is naturally defined as

$$\langle f \rangle \equiv \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{m=1}^M \langle N_m \rangle f_m, \quad (2.1.6)$$

so that using Equation (2.1.3) we get

**Expectation value via probabilities:**

$$\langle f \rangle = \sum_{m=1}^M W_m f_m. \quad (2.1.7)$$

Notice that Equation (2.1.3) may be considered as the particular form of this general result, when all  $f_m = 1$ .

Next, the spectrum of possible experimental outcomes is frequently continuous for all practical purposes. (Think, for example, about the set of positions of the marks left by bullets fired into a target from afar.) The above formulas may be readily generalized to this case; let us start from the simplest situation when all different outcomes may be described by just one continuous scalar variable  $q$  – which replaces the discrete index  $m$  in Eqs. (2.1.1)-(2.1.7). The basic relation for this case is the self-evident fact that the probability  $dW$  of having an outcome within a small interval  $dq$  near some point  $q$  is proportional to the magnitude of that interval:

$$dW = w(q)dq, \quad (2.1.8)$$

where  $w(q)$  is some function of  $q$ , which does not depend on  $dq$ . This function is called *probability density*. Now all the above formulas may be recast by replacing the probabilities  $W_m$  with the products (2.1.8), and the summation over  $m$ , with the integration over  $q$ . In particular, instead of Equation (2.1.4) the normalization condition now becomes

$$\int w(q)dq = 1, \quad (2.1.9)$$

where the integration should be extended over the whole range of possible values of  $q$ . Similarly, instead of the discrete values  $f_m$  participating in Equation (2.1.5), it is natural to consider a function  $f(q)$ . Then instead of Equation (2.1.7), the expectation value of the function may be calculated as

**Expectation value via probability density:**

$$\langle f \rangle = \int w(q)f(q)dq. \quad (2.1.10)$$

It is also straightforward to generalize these formulas to the case of more variables. For example, the state of a classical particle with three degrees of freedom may be fully described by the probability density  $w$  defined in the 6D space of its generalized radius-vector  $\mathbf{q}$  and momentum  $\mathbf{p}$ . As a result, the expectation value of a function of these variables may be expressed as a 6D integral

$$\langle f \rangle = \int w(\mathbf{q}, \mathbf{p})f(\mathbf{q}, \mathbf{p})d^3q d^3p. \quad (2.1.11)$$

Some systems considered in this course consist of components whose quantum properties cannot be ignored, so let us discuss how  $\langle f \rangle$  should be calculated in this case. If by  $f_m$  we mean measurement results, then Equation (2.1.7) (and its generalizations) remains valid, but since these numbers themselves may be affected by the intrinsic quantum-mechanical uncertainty, it may make sense to have a bit deeper look into this situation. Quantum mechanics tells us<sup>4</sup> that the most general expression for the expectation value of an observable  $f$  in a certain ensemble of macroscopically similar systems is

$$\langle f \rangle = \sum_{m,m'} W_{mm'} f_{m'm} \equiv \text{Tr}(Wf). \quad (2.1.12)$$

Here  $f_{mm'}$  are the matrix elements of the quantum-mechanical operator  $\hat{f}$  corresponding to the observable  $f$ , in a full basis of orthonormal states  $m$ ,

$$f_{mm'} = \langle m | \hat{f} | m' \rangle, \quad (2.1.13)$$

while the coefficients  $W_{mm'}$  are the elements of the so-called *density matrix*  $W$ , which represents, in the same basis, the *density operator*  $\hat{W}$  describing properties of this ensemble. Equation (2.1.12) is evidently more general than Equation (2.1.7), and is reduced to it only if the density matrix is diagonal:

$$W_{mm'} = W_m \delta_{mm'} \quad (2.1.14)$$

(where  $\delta_{mm'}$  is the Kronecker symbol), when the diagonal elements  $W_m$  play the role of probabilities of the corresponding states.

Thus formally, the largest difference between the quantum and classical description is the presence, in Equation (2.1.12), of the off-diagonal elements of the density matrix. They have the largest values in the *pure* (also called “coherent”) ensemble, in which the

state of the system may be described with *state vectors*, e.g., the *ket-vector*

$$|\alpha\rangle = \sum_m \alpha_m |m\rangle \quad (2.1.15)$$

where  $\alpha_m$  are some (generally, complex) coefficients. In this case, the density matrix elements are merely

$$W_{mm'} = \alpha_m^* \alpha_{m'}, \quad (2.1.16)$$

so that the off-diagonal elements are of the same order as the diagonal elements. For example, in the very important particular case of a *two-level system*, the pure-state density matrix is

$$W = \begin{pmatrix} \alpha_1^* \alpha_1 & \alpha_1^* \alpha_2 \\ \alpha_2^* \alpha_1 & \alpha_2^* \alpha_2 \end{pmatrix}, \quad (2.1.17)$$

so that the product of its off-diagonal components is as large as that of the diagonal components.

In the most important basis of stationary states, i.e. the eigenstates of the system's time independent Hamiltonian, the coefficients  $\alpha_m$  oscillate in time as<sup>5</sup>

$$\alpha_m(t) = \alpha_m(0) \exp\left\{-i \frac{E_m}{\hbar} t\right\} \equiv |\alpha_m| \exp\left\{-i \frac{E_m}{\hbar} t + i \varphi_m\right\}, \quad (2.1.18)$$

where  $E_m$  are the corresponding eigenenergies, and  $\varphi_m$  are constant phase shifts. This means that while the diagonal terms of the density matrix (2.1.16) remain constant, its off-diagonal components are oscillating functions of time:

$$W_{mm'} = \alpha_{m'}^* \alpha_m = |\alpha_{m'}| |\alpha_m| \exp\left\{i \frac{E_m - E_{m'}}{\hbar} t\right\} \exp\{i (\varphi_{m'} - \varphi_m)\} \quad (2.1.19)$$

Due to the extreme smallness of the Planck constant (on the human scale of things), minuscule random perturbations of eigenenergies are equivalent to substantial random changes of the phase multipliers, so that the time average of any off-diagonal matrix element tends to zero. Moreover, even if our statistical ensemble consists of systems with exactly the same  $E_m$ , but different values  $\varphi_m$  (which are typically hard to control at the initial preparation of the system), the average values of all  $W_{mm'}$  (with  $m \neq m'$ ) vanish again.

This is why, besides some very special cases, typical statistical ensembles of quantum particles are far from being pure, and in most cases (certainly including the thermodynamic equilibrium), a good approximation for their description is given by the opposite limit of the so-called *classical mixture*, in which all off-diagonal matrix elements of the density matrix equal zero, and its diagonal elements  $W_{mm}$  are merely the probabilities  $W_m$  of the corresponding eigenstates. In this case, for the observables compatible with energy, Equation (2.1.12) is reduced to Equation (2.1.7), with  $f_m$  being the eigenvalues of the variable  $f$ , so that we may base our further discussion on this key relation and its continuous extensions (2.1.10)-(2.1.11).

---

This page titled [2.1: Statistical ensemble and probability](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Konstantin K. Likharev](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.