Time-Dependent Statistical Mechanics
3. Elementary Probability Theory

H.C.A.

September 24, 2009

In these notes, we will review some ideas of elementary probability theory that we need in this course. We will not discuss this in class, but you should read these notes before the second lecture, if possible. We will make extensive use of these ideas and this notation in the course.

1 Elementary probability theory

The contents of this section are some elementary concepts from probability theory, stated in the language that is appropriate for the way in which we use these concepts in statistical mechanics. These notes are not intended as a textbook in probability. I am assuming that you already have some background in probability theory from other chemistry, physics, and math courses. If any of these ideas is confusing or unfamiliar, then please discuss it with me.

Random variables. Associated with any physical system of interest is a set of random variables. They include things like:

- The x coordinate of the first atom of the system at the time t_0.
- The number of atoms in the system.
- The y component of the velocity of the second atom of the system at time t_1, where $t_1 > t_0$.

- The net distance traveled by the fourth atom between times t_0 and t_1.
- The vector force on the fifth atom at time t_2.

If any of these quantities is not precisely controlled by the experimental protocol, then in general different values will be found each time an experiment is performed. For our purposes, a *random variable* is a well defined physical quantity whose value is not determined by the way in which the experiment is performed but which in general takes on different values each time the experiment is performed.[1]

Distribution function for a random variable. Let X be a random variable. It distribution function, $P_X(x)$, is defined so that $P_X(x)dx$ = the probability that X has a value between x and $x + dx$. This function has a number of properties.

- $P_X(x)dx$ is also the fraction of the experiments for which X has a value between x and $x + dx$.
- $P_X(x) \geq 0$.
- $$\int_a^b dx\, P_X(x) = \text{the probability that } X \text{ has a value between } a \text{ and } b.$$
- The normalization condition.
$$\int_{-\infty}^{\infty} dx\, P_X(x) = 1$$

Expectation value of a random variable. If X is a random variable, its expectation value $\langle X \rangle$ is defined as

$$\langle X \rangle = \int_{-\infty}^{\infty} dx\, x P_X(x)$$

It is equal to the average value of the result of many experimental measurements of X in independent experiments.

[1]In this discussion we shall treat all random variables as if they were real numbers. The discussion could easily be generalized to the case of random variables that are arrays (or vectors) of real numbers or that take on discrete values or that are complex.

Expectation value of a function of a random variable. Suppose Z is a random variable, but it is a function of another random variable X. For example, X might be the x component of momentum of atom 1 and Z might be $X^2/2m$, where m is the mass of the atom. Thus Z has the physical meaning of the kinetic energy associated with the x momentum of atom 1. Both X and Z are random variables, but the first completely determines the second. Under these circumstances, the expectation value of Z can be calculated from the distribution function of X.

Suppose $Z = Z(X)$, i.e. Z is some well defined function of X. Then

$$\langle Z \rangle = \int_{-\infty}^{\infty} dx\, P_X(x) Z(x)$$

E.g. for the example in the previous paragraph,

$$\langle Z \rangle = \langle X^2/2m \rangle = \int_{-\infty}^{\infty} dx\, P_X(x) x^2/2m$$

Joint distribution functions for two random variables. Suppose X and Y are two random variables. We can imagine measuring both quantities for the same experiment. The *joint distribution function of two variables*, $P_{XY}(x, y)$, is defined so that $P_{XY}(x, y)dxdy =$ the probability that X has a value between x and $x + dx$ and that Y has a value between y and $y + dy$. This function has a number of properties.

- $P_{XY}(x, y)dxdy$ is also the fraction of the experiments for which X has a value between x and $x + dx$ and Y has a value between y and $y + dy$.
- $P_{XY}(x, y) \geq 0$.
- $$\int_a^b dx \int_c^d dy\, P_{XY}(x, y) =$$
  the probability that X has a value between a and b and that Y has a value between c and d.
- $$\int_{-\infty}^{\infty} dy\, P_{XY}(x, y) = P_X(x)$$
  $$\int_{-\infty}^{\infty} dx\, P_{XY}(x, y) = P_Y(y)$$
  This follows from the previous result and from the definition of $P_X(x)$.

- $$\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, P_{XY}(x,y) = 1$$

Some elementary properties of the expectation value.

- The expectation value of a sum is the sum of the expectation values.

$$\langle X + Y \rangle = \langle X \rangle + \langle Y \rangle$$

- If a is a constant and X is a random variable, $\langle aX \rangle = a\langle X \rangle$.

  Note, however, that it is not in general true that

$$\langle XY \rangle = \langle X \rangle \langle Y \rangle$$

Statistical independence of random variables. Two random variables X and Y are *statistically independent* if $P_{XY}(x,y) = P_X(x)P_Y(y)$.

Statistical independence of two random variables has a number of consequences.

- On an intuitive level, it means that if I tell you the value of X that is measured in an experiment, this gives you no information that might help you to guess the value of Y.

- It means that

$$\langle XY \rangle = \langle X \rangle \langle Y \rangle$$

  The proof is straightforward.

$$\begin{aligned}\langle XY \rangle &= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, xyP_{XY}(x,y) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, xyP_X(x)P_Y(y) \\ &= \int_{-\infty}^{\infty} dx\, xP_X(x) \int_{-\infty}^{\infty} dy\, yP_Y(y) = \langle X \rangle \langle Y \rangle\end{aligned}$$

Note that if two variables X and Y do *not* satisfy this equation, then then can *not* be statistically independent. However, if they do satisfy this equation, they might or might not be statistically independent.

Three or more random variables. Similarly, we can define distribution functions for three or more random variables.