Chapter 3

Gaussian Distributions

3.1 Introduction

The simplest and most symmetrical problem in statistical mechanics uses a variable, like the classical momentum, which runs over all values in the interval between minus infinity and infinity. In this chapter, we use the symbol X for such a variable and $X_j (j = 1, 2, \ldots, N)$ for N such variables. We might expect that such a variable has so much symmetry that it has the potential for producing situations which are both particularly tractable and also particularly robust. By robust one means that the basic problems or their solutions remain unchanged under a wide class of transformations. We start here by considering the probabilistic behavior with one Gaussian variable, X, and move on to the consideration of many such variables.

In one sense the discussion of this chapter is a natural outgrowth of the last one. There we were interested in the comparison between the structure of problems of the statistical mechanics of one particle and the ones with many non-interacting particles. With Gaussian variables, if the probability distribution is proportional to the exponential of a quadratic form in X (or the X_j's) then the problem may once more be solved exactly as if it were a problem of non-interacting systems.

3.2 One Variable

A Gaussian probability distribution for a single random variable X is one in which the probability of finding X to lie between x and $x + dx$ has the form[1]:

$$\frac{\exp[-\beta(x - \langle X \rangle)^2/2]}{\sqrt{2\pi/\beta}} dx. \tag{3.1}$$

This distribution contains two parameters, $\langle X \rangle$, which measures the center, or mean of the distribution and β, which measures its squared width. This squared width is called the

[1]The word Gaussian is also used to describe an integral in which the integrand is the exponential of a quadratic form, i.e. $\int_{-\infty}^{\infty} dx \exp(-ax^2 + bx + c)$.

variance. As we have already seen, β is the inverse of the variance in the distribution, so that

$$\beta^{-1} = \langle (X - \langle X \rangle)^2 \rangle \,. \tag{3.2}$$

Any probability distribution in one variable, x, which varies over the entire real line and which is of the form of an exponential of a function quadratic in x

$$\rho \sim \exp(-ax^2 + bx + c)$$

may be cast into the structure shown on the right hand side of Eq. (3.1). (See Problem (3.1).)

Our work in describing statistical mechanics has relied quite heavily upon understanding the behavior of exponential functions of the basic variables. We continue along this track by analyzing the so-called *generating function*, $\langle \exp(iqX) \rangle$, with i being the square root of -1 and q being real. In this context, a generating function is the name for an average which is used to generate (i.e. *compute*) other averages. For the Gaussian case, our generating function is the integral

$$\langle \exp(iqX) \rangle = \int_{-\infty}^{\infty} dx \frac{\exp(-\beta(x - \langle X \rangle)^2/2}{\sqrt{2\pi/\beta}} \exp(iqx)$$

which is easily evaluated to get the result:

> For any Gaussian probability distribution with variance β^{-1}, the average of an exponential takes the form
>
> $$\langle \exp(iqX) \rangle = \exp\left(iq\langle X \rangle - \frac{q^2}{2\beta}\right) . \tag{3.3}$$
>
> Conversely, if some variable, X, has a probability distribution for which
>
> $$\langle \exp(iqX) \rangle = \exp\left[iqa - \frac{q^2}{2\beta}\right] \tag{3.4}$$
>
> for all values of q, then the probability distribution for X must be Gaussian and a and β respectively have the interpretation of the mean and the inverse variance of the Gaussian.

These two statements look like very special results. Why should they be interesting or important?

One reason is that the Gaussian distribution exhibits a remarkable tenacity or stability, or what we call a peculiar robustness. Imagine that the variable X is of the form $X = aY + b$ where a and b are constants and Y is a Gaussian variable. Equation (3.3) then has the consequence that X is also a Gaussian random variable. Thus a linear function of a Gaussian random variable is also a Gaussian random variable. Next consider an X of the form

$$X = aY + bZ \,, \tag{3.5a}$$

where a and b are constants and Y and Z are independent Gaussian random variables. By independent it is meant that the fluctuations in the two variables are uncorrelated or *independent*. The mathematical definition of independence use the joint probability $\rho(z,y)dz\,dy$. This expression gives the probability for finding Z between z and $z+dz$ and, at the same time, finding Y between y and $y+dy$. Independence is that statement that the joint probability distribution $\rho(z,y)$ is a product of the individual probability distributions for Z and Y.

Thus it follows at once from Eq. (3.3) and the meaning of the word independent that X is also a Gaussian random variable. If you add up many independent Gaussian variables you end up with a new Gaussian variable. Thus, if

$$X = \sum_{j=1}^{N} a_j Y_j\,, \tag{3.5b}$$

where the a_j are constants and the Y_j are independent Gaussian random variables then the resulting X is certainly a Gaussian random variable. Consequently, a sum of very many Gaussian physical effects (represented say by the Y's) is a net Gaussian effect.

3.3 Many Gaussian Variables

To describe the effect of many Gaussian Variables, we generalize Eqs. (3.5a) and (3.5b): $X_1, X_2, \ldots, X_j, \ldots, X_N$ are said to be a set of *Gaussian variables* if *the expectation value of an exponential formed from a linear combination of these variables is an exponential of a quadratic form in the coefficients*. The words are a mouthful, the equation is simple. For any vector $\mathbf{q}$ with components $q_j (j = 1, 2, \ldots, N)$ we have

$$\left\langle \exp\left(i \sum_{j=1}^{N} q_j X_j \right) \right\rangle = \exp\left(i \sum_{j=1}^{N} q_j \langle X_j \rangle - \sum_{j,k=1}^{N} q_j G_{jk} \frac{q_k}{2} \right). \tag{3.6}$$

Thus, the quadratic form defines this many-variable Gaussian.[2] Naturally, we shall usually write Eq. (3.6) in matrix-vector notation as

$$\langle \exp(i\mathbf{q} \cdot \mathbf{X}) \rangle = \exp\left(i\mathbf{q} \cdot \langle \mathbf{X} \rangle - \mathbf{q} \cdot G \cdot \frac{\mathbf{q}}{2} \right). \tag{3.7}$$

Of course $\langle \mathbf{X} \rangle$ means the average of the vector $\mathbf{X}$ and G is related to a variance or correlation matrix for these variables. In fact, a second order power series expansion (in $\mathbf{q}$) of Eq. (3.7) implies:

$$\langle (X_j - \langle X_j \rangle)(X_k - \langle X_k \rangle) \rangle = G_{jk}\,. \tag{3.8}$$

[2] Analytic continuation enables us to extend this results to complex values of the q's.

For the definition (3.6) to make sense, we require that G be a positive definite symmetric matrix.[3] Often, G is called a green function.

What probability distribution could give rise to averages of the form (3.6)? The reader will not be surprised to hear that the defining probability distribution must be an exponential of a quadratic form. Specifically, I ask you to prove in a homework exercise that if

$$\rho(x_1, x_2, \dots, x_N) dx_1\, dx_2 \cdots x_N$$

is defined to be the probability that X_1 have a value between x_1 and $x_1 + dx_1$ while X_2 has a value between x_2 and $x_2 + dx_2$, ... then from (3.6) it follows that:

$$\rho(x_1, x_2, \dots, x_N) = (2\pi)^{-N/2} (\det C)^{1/2} e^{-[(x-\langle X\rangle)C(x-\langle X\rangle)]/2} \tag{3.9}$$

Here C is a symmetric matrix which is the inverse of G, namely

$$\sum_{m=1}^{N} C_{jm} G_{mk} = \sum_{m=1}^{N} G_{jm} C_{mk} = \delta_{j,k} \tag{3.10}$$

or in matrix form

$$CG = GC = 1\,. \tag{3.11}$$

The determinant of C has a value which is the product of all the eigenvalues of C.

The key results in obtaining the solution (3.10) are Eqs. (3.6) and (3.9). That solution makes the many-variable Gaussian case look just like the one-variable case. Imagine doing a change of variables from the X's to the appropriate linear combination of X's which serves to diagonalize the matrix C. Then G is simultaneously diagonalized. After this replacement the probability density is an exponential of a sum of terms, each one quadratic in one of the X-variables. Then, Eq. (3.6) gives the total probability as a product of probabilities for the individual linear combinations. Thereby the entire problem is reduced to a combination of subproblems for the individual and non-correlated linear combinations.

We have defined Gaussian correlations by using a generating function approach — and G appears in the solution for the generating function. Conversely, Eq. (3.9) shows that the matrix inverse of G, called C, appears in the basic problem definition. Thus to go from problem to solution, one need only invert a matrix.

[3] Any anti-symmetric part of G will contribute nothing to the right hand side of Eq. (3.7). Hence we do not lose any generality by demanding that G is a symmetric matrix. Any symmetric real matrix is Hermitian. One can describe it by listing its eigenvalues, all of which must be real. If they are all positive then the matrix is said to be positive definite. However, if one eigenvalue is negative, then one can set up a Y as a linear combination of X_j corresponding to the eigen direction. A negative eigenvalue of G corresponds to a negative value of $\langle (Y - \langle Y\rangle)^2\rangle$. This negativity is quite impossible for real Y. Consequently, only positive eigenvalues are possible in a formulation based upon Eq. (3.8).

An alternate formulation is possible, based upon Gaussian integrals. In that case the positivity is a condition for the convergence of the integrals.