Reading assignment. Schroeder, section **2.5**.

1 Large systems

Following up on the last two themes, the observation of increased number of associated microstates (or multiplicity) of macrostates having more even partitioning of macroscopic parameters (only energy was considered) between subsystems, and the ability to use Stirling's approximation for factorials of large numbers, we'll take an explicit look at the narrowing of the probability distribution for macrostates next. We could look at the probability distribution of macrostates of a pair of interacting subsystems, like we did before, but the manipulations and the interpretation will be a bit simpler if we confine ourselves to a single system. We don't want to isolate that system completely, though, lest we be stuck with only a single accessible macrostate.

So, we'll work with a single two-state paramagnet in thermal contact with a second, unspecified system. This permits the two systems to exchange energy, but we'll look only at the states of the paramagnet, rather than both together.

Recall that the energy in the presence of an external magnetic field is proportional to the number of magnetic moments that are antialigned with respect to the field, if we define the ground state to have zero energy. Each macrostate is characterized by the number of flipped moments, and the corresponding number of microstates, the multiplicity, is given by the binomial coefficient

$$\Omega(N,k) = \binom{N}{k} = \frac{N!}{k!(N-k)!}\,, \tag{1}$$

k being the number of flipped moments.

Now suppose that the interaction with the other system causes some energy to be dumped into our paramagnet, flipping some of the moments. Further, suppose the result of this is that the moments of the paramagnet individually have a probability α of being flipped. Then the probability distribution for the number of flipped moments is just

$$P(k) = \binom{N}{k}\alpha^k(1-\alpha)^{N-k}\,. \tag{2}$$

The idea there is that the flipping or lack thereof of individual moments is independent of all the others, so the compound probability of having a particular set of k moments flipped and the other $N-k$ moments not flipped is the product of the individual probabilities, $\alpha^k(1-\alpha)^{N-k}$. But there are $\binom{N}{k}$ ways of choosing the k moments to flip, leading to the probability distribution above. That distribution is a famous one, the *binomial distribution.*

Before proceeding further, let's check the normalization:

$$
\begin{aligned}
\sum_{k=0}^{N} P(k) &= \sum_{k=0}^{N} \binom{N}{k} \alpha^k (1-\alpha)^{N-k} \\
&= [\alpha + (1-\alpha)]^N \\
&= 1\,.
\end{aligned}
\tag{3}
$$

The equality on the second line follows from the familiar expansion of a binomial:

$$
(a+b)^N = \sum_{k=0}^{N} \binom{N}{k} a^k b^{N-k}\,. \tag{4}
$$

Now the single-moment probability α is closely related to the mean number of flipped moments, a relationship we can determine by calculating the average value of k. This is done by the usual weighted average with the probability distribution $P(k)$ providing the weights:

$$
\begin{aligned}
\langle k \rangle &= \sum_{k=0}^{N} k P(k) \\
&= \sum_{k=1}^{N} k \frac{N!}{k!(N-k)!} \alpha^k \beta^{N-k} \qquad (\text{let } \beta = 1-\alpha) \\
&= \sum_{k=1}^{N} \frac{N!}{(k-1)!(N-k)!} \alpha^k \beta^{N-k}\,.
\end{aligned}
\tag{5}
$$

This is somewhat similar to the normalization sum seen above, but the lower limit of the sum is 1, rather than 0, and the first factorial in the denominator is of $k-1$, instead of k. Taking that as a hint, let's try shifting the indices to see if we can recover something like the normalization sum as a portion of the result. To that end, let

$$
l = k-1 \quad \text{and} \quad M = N-1\,, \tag{6}
$$

which leads to

$$
\begin{aligned}
\langle k \rangle &= \sum_{l=0}^{M} \frac{(M+1)!}{l!(M+1-l-1)!} \alpha^{l+1} \beta^{M+1-l-1} \\
&= (M+1)\alpha \sum_{l=0}^{M} \frac{M!}{l!(M-l)!} \alpha^l \beta^{M-l} \\
&= N\alpha\,.
\end{aligned}
\tag{7}
$$

The last step follows from recognition of the sum as the Mth power of $\alpha + \beta = 1$, together with the identification $M + 1 = N$. Thus, α is just $\langle k \rangle / N$, the fractional mean moment flip.

To characterize the width of the probability distribution, we could directly calculate the standard deviation by a method similar to that used to find the mean. Instead, we'll look at what happens to the distribution in the large-N limit, which is our primary interest anyway. We'll use Stirling's approximation to handle the factorials of large numbers:

$$N! \sim N^N e^{-N} \sqrt{2\pi N}\,. \tag{8}$$

The large-N behavior of $P(k)$ is then

$$\begin{aligned} P(k) &= \frac{N!}{k!(N-k)!}\alpha^k \beta^{N-k} \\ &\sim \frac{N^N e^{-N}\sqrt{2\pi N}\alpha^k\beta^{N-k}}{k^k e^{-k}\sqrt{2\pi k}(N-k)^{N-k}e^{-N+k}\sqrt{2\pi(N-k)}} \\ &= \left(\frac{N\alpha}{k}\right)^k \left(\frac{N\beta}{N-k}\right)^{N-k} \sqrt{\frac{N}{2\pi k(N-k)}}\,. \end{aligned} \tag{9}$$

Now, we'll want to see how the probability falls off away from the mean value of k, so we'll change to a variable measured relative to that:

$$x = k - N\alpha\,, \tag{10}$$

which has the probability distribution

$$P(x) = \underbrace{\left(\frac{N\alpha}{N\alpha + x}\right)^{N\alpha+x} \left(\frac{N\beta}{N\beta - x}\right)^{N\beta-x}}_{\text{call this } f(x)} \underbrace{\sqrt{\frac{N}{2\pi k(N-k)}}}_{\text{ignore this}} \tag{11}$$

We can make further progress by working with the logarithm of f, which will allow us to make use of the expansion of $\ln(1+\epsilon)$:

$$\begin{aligned} \ln f(x) &= (N\alpha + x)\left[\ln(N\alpha) - \ln(N\alpha + x)\right] \\ &\quad + (N\beta - x)\left[\ln(N\beta) - \ln(N\beta - x)\right] \\ &= (N\alpha + x)\left[\ln(N\alpha) - \ln(N\alpha) - \ln\left(1 + \frac{x}{N\alpha}\right)\right] \\ &\quad + (N\beta - x)\left[\ln(N\beta) - \ln(N\beta) - \ln\left(1 - \frac{x}{N\beta}\right)\right] \end{aligned} \tag{12}$$

Since we're mainly interested in $x \ll N\alpha$ or $N\beta$, we'll approximate the logs by truncated series expansions:

$$\ln(1+\epsilon) = \epsilon - \frac{\epsilon^2}{2} + \frac{\epsilon^3}{3} - \ldots \tag{13}$$

keeping two terms. This gives

$$
\begin{aligned}
\ln f(x) &\approx (N\alpha + x)\left[-\frac{x}{N\alpha} + \frac{1}{2}\left(\frac{x}{N\alpha}\right)^2\right] \\
&\quad + (N\beta - x)\left[\frac{x}{N\beta} + \frac{1}{2}\left(\frac{x}{N\beta}\right)^2\right] \\
&\approx -x - \frac{x^2}{N\alpha} + \frac{x^2}{2N\alpha} + x - \frac{x^2}{N\beta} + \frac{x^2}{2N\beta} \\
&= -\frac{x^2}{2N}\left(\frac{1}{\alpha} + \frac{1}{\beta}\right) \\
&= -\frac{x^2}{2N\alpha\beta}\,.
\end{aligned}
\tag{14}
$$

Exponentiating this gives

$$
f(x) \sim e^{-x^2/2N\alpha\beta}\,, \tag{15}
$$

which is just a Gaussian.

The probability distribution is proportional to this, but the square root we're ignoring is not the correct normalization factor for the Gaussian distribution. We won't need the correct normalization, though.

We want to measure the deviation from the mean ($x = 0$) in units of the full width of the original binomial distribution, so we'll define

$$
y = \frac{x}{N}\,, \tag{16}
$$

which would, in principle, allow us to plot Gaussians for different numbers N of magnetic moments on the same graph in order to compare their widths. In terms of the scaled variable y, our function is now

$$
f(y) = e^{-N^2y^2/2N\alpha\beta} = e^{-Ny^2/2\alpha\beta}\,. \tag{17}
$$

Away from $y = 0$ this reaches $1/e$ times the maximum value at

$$
f(y_e) = \frac{1}{e}\,, \tag{18}
$$

which yields

$$
\frac{Ny_e^2}{2\alpha\beta} = 1\,, \tag{19}
$$

or

$$
y_e = \sqrt{\frac{2\alpha\beta}{N}}\,. \tag{20}
$$

Since y_e is proportional to $1/\sqrt{N}$, if $N = 10^{20}$, then the width is of order 10^{-10} of the width of the plot, a rather skinny distribution, indeed.

This means that macroscopic fluctuations away from the mean number of flipped moments are exceedingly unlikely—virtually nonexistent in macroscopic systems. The most probable macrostate is the only one with any significant probability of being observed.

Formally, one sometimes wants to speak of the actual limit of infinite system size, in which fluctuations vanish completely—then the probability distribution becomes a δ function. That limit is called the *thermodynamic limit*, but the term is also often used loosely to refer to any macroscopic system.

HW Problem. Schroeder problem 2.22, p. 66.