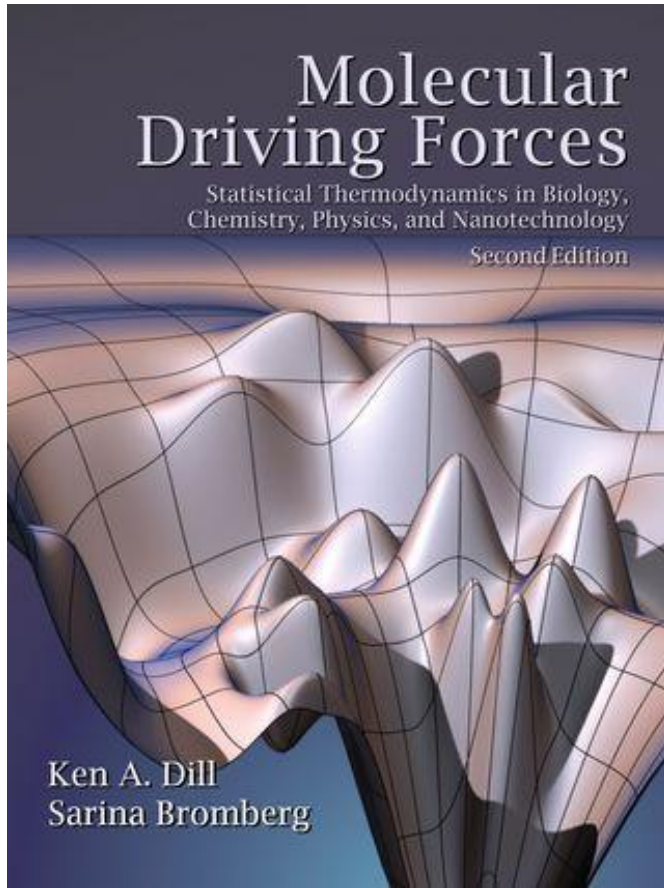# Statistical thermodynamics

Mark HUI

yyhui@pub.iams.sinica.edu.tw

Lab.: Rm 510, Off.: Rm 317

Reference

- Atkins' 8th edition
- Chapter 16, 17

Statistical thermodynamics

What Statistical thermodynamics studies?

- Solid melts .
- Crystal grows.
- Chemicals react and take up energy or give it off.
- Protein catalyze biological reactions
- What forces drive these processes?
- Statistical thermodynamics model the molecular forces to drive these processes.

Entropy and energy

- We can measure the density, temperature, pressure, capacity, molecular radius... to understand this process.

- But, we cannot simply predict the tendency and equilibria of system.

- To predict the equilibria, we must step into a different world. Energy, entropy, enthalpy and free energy.

- For instance, water boils at 100^OC. But measure the density of water at 98 ^{o}C, we do not know the sudden change at 100^oC.

- To predict the density change, we need to know the driving force, the entropies and the energies.

Entropy

- Describes the tendency of matter towards disorder.

- Entropy explains heat flows from hot object to cold object.

- How protein molecules tangle together in some disease state.

Probability

- $P_A = n_A / N$
- N=total number of possible outcome
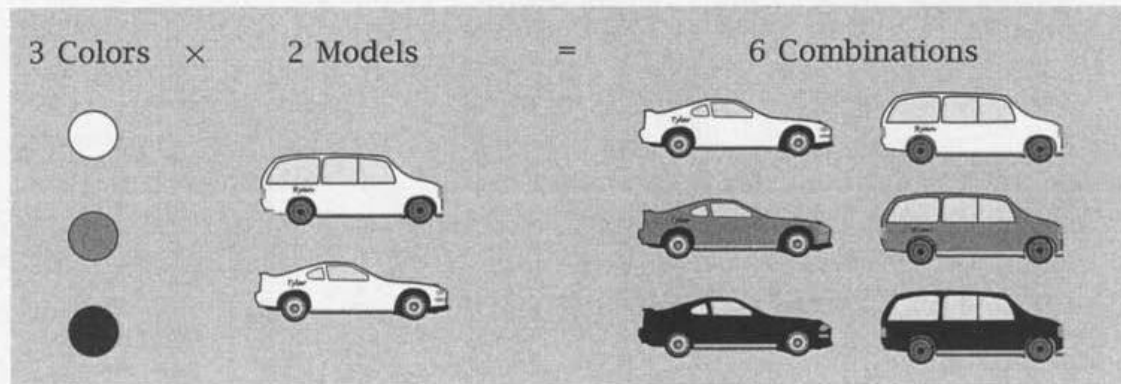- n_A= the outcomes fall into catalog A

Figure 1.1 If there are three car colors for each of two car models, there are six different combinations of color and model, so the multiplicity is six.

Multiplication rule

If outcome A, B are independent, then the probability of observing A and B is

P(A AND B)=$P_A P_B$

- What is the probability of getting 4 heads on 4 successive flips of an unbiased coin?

- What is the probability of getting one heads, then one tail, then two heads on 4 successive coin flips? (In this specific order)

Examples for Multiplication rule

- In previous example, what is the probability of the specific sequence of 4 coin flips, HTHH?

- What is the probability of obtaining 3H and 1T in any order?

- HHHT, HHTH, HTHH, THHH

- 4 out of 16 possible outcomes

Permutations of ordered sequences

- How many permutations of ordered sequences of the letter w, x, y and z?
- For instance, wxyz, wyzx
- 1st letter can be any 1 of the character
- 2nd letter can be one of the 3 letters
- ...
- W=4x3x2x1

Permutations of ordered sequences

- In general, a sequence of N distinguishable objects, the number of different permutations
- *W=N!*

Counting sequences of distinguishable and indistinguishable objects

- How many sequence can we arrange (H, A_1, A_2)?
- (HA_1A_2) (A_1HA_2) (A_1A_2H)

(HA_2A_1) (A_2HA_1) (A_2A_1H)

- $W = 3!$

If A_1 and A_2 are indistinguishable,

$(A_2HA_1) = (A_2A_1H)$

- *$W = 3!/2!$*

Number of permutations

In general, for a collection of N objects with t categories, of which n_i objects in each category are *indistinguishable* from one another, but distinguishable from the objects in the other $t - 1$ categories, the number of permutations W is

$$W = \frac{N!}{n_1! n_2! \cdots n_t!}. \tag{1.18}$$

When there are only two categories (success/failure, or heads/tails, ...), $t = 2$ so $W(n, N)$, the number of sequences with n successes out of N trials, is

$$W(n, N) = \binom{N}{n} = \frac{N!}{n!(N - n)!}, \tag{1.19}$$

Number of permutations from indistinguishable objects

EXAMPLE 1.15 Counting sequences of coin flips and die rolls. You flip a coin $N = 4$ times. How many different sequences have three heads? According to Equation (1.19),

$$W(n_H, N) = \frac{N!}{n_H! n_T!} = \frac{4!}{3! 1!} = 4.$$

They are $THHH, HTHH, HHTH$, and $HHHT$. How many different sequences have two heads?

$$W(2, 4) = \frac{4!}{2! 2!} = 6.$$

They are $TTHH, HHTT, THTH, HTHT, THHT$, and $HTTH$.

You flip a coin one hundred and seventeen times. How many different sequences have thirty-six heads?

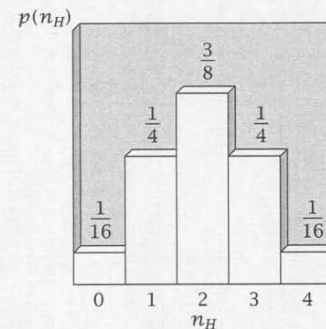$$W(36, 117) = \frac{117!}{36! 81!} \approx 1.84 \times 10^{30}.$$

Figure 1.5 The probability distribution for the numbers of heads in four coin flips in Example 1.18.

13

Number of permutations from indistinguishable objects

You roll a die fifteen times. How many different sequences have three 1's, one 2, one 3, five 4's, two 5's, and three 6's? According to Equation (1.18),

$$W = \frac{15!}{3!\,1!\,1!\,5!\,2!\,3!} = 151,351,200.$$

Bose-Einstein condensates

(a) Balls in Boxes

EXAMPLE 1.17 Bose–Einstein statistics. How many ways can n indistinguishable particles be put into M boxes, with any number of particles per box? This type of counting is needed to predict the properties of particles called *bosons*, such as photons and He^4 atoms. Bose–Einstein statistics counts the ways that n particles can be distributed in M different energy levels, when several particles can occupy the same quantum mechanical energy levels. For now, our interest is not in the physics, but just in the counting problem. Figure 1.3 shows that one way to count the number of arrangements is to think of the system as a linear array of n particles interspersed with $M-1$ movable walls that partition the system into M boxes (spaces between walls). There are $M+n-1$ objects, counting walls plus particles. The n particles are indistinguishable from each other. The $M-1$ walls are indistinguishable from the other walls. Because the walls are distinguishable from the particles, the number of arrangements is

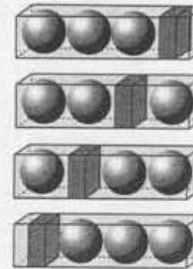$$W(n,M) = \frac{(M+n-1)!}{(M-1)!n!}. \tag{1.20}$$
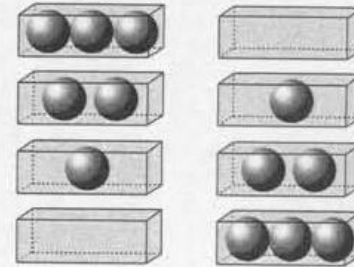
(b) Moveable Walls

Figure 1.3 Three Bose-Einstein particles in two boxes for Example 1.17: (a) There are four ways to partition $n = 3$ balls into $M = 2$ boxes when each box can hold any number of balls. (b) There are also four ways to partition three balls and one movable wall.

What are Extremum Principles?

- The forces on atoms and molecules can be described in terms of two tendency, energy and entropy.

- Molecules react, change conformations, bind, and undergo chemical and physical changes, in the way to reach the minimum energy and maximum entropy.

- We can predict the tendencies of the matters by computing the minima or the maxima of a certain mathematical function.

- Extremum principles or variational principles.

Mechanical equilibrium

- A ball rolls down until it reaches the bottom and stay at the bottom, equilibrium position.
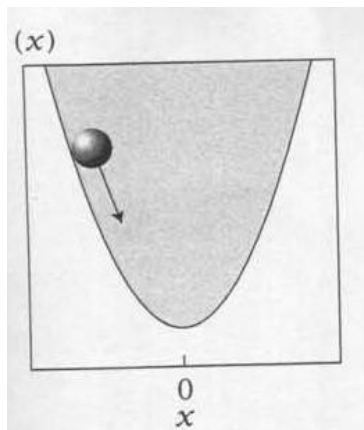
(x)

0
x

igure 2.1   The equilibrium
osition $x^* = 0$ for a ball in
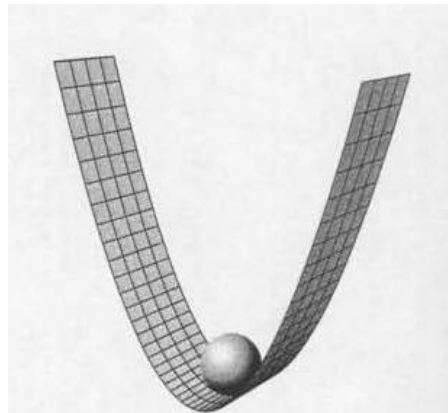quadratic valley has the
ninimum potential energy
$'(x^*)$.
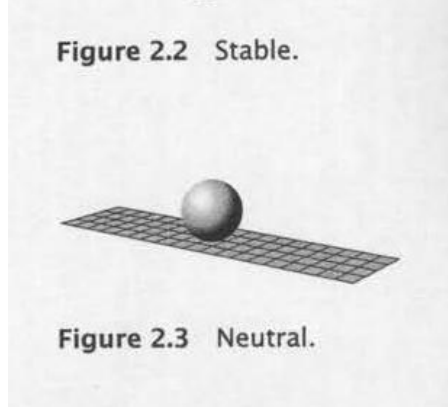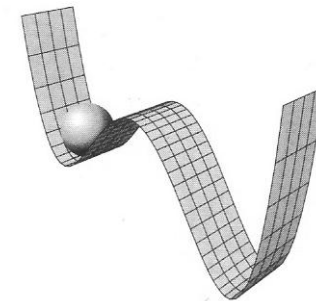
Figure 2.2   Stable.

Figure 2.3   Neutral.

Figure 2.4   Metastable.
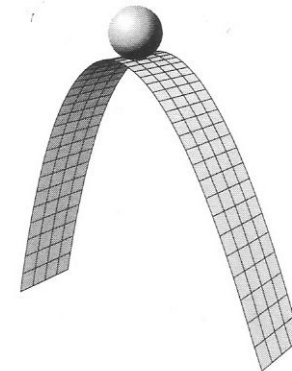
Figure 2.5   Unstable.

Extremeum principle: Maximizing multiplicity predicts the most probable outcomes.

- For instance, flipping a coin 4 times.
- Which composition is more probable?
- 3HIT or 4H

Table 2.1 $N = 4$

n	W	$\ln W$
4	$\dfrac{4!}{0!4!} = 1$	0
3	$\dfrac{4!}{1!3!} = 4$	1.386
2	$\dfrac{4!}{2!2!} = 6$	1.792
1	$\dfrac{4!}{3!1!} = 4$	1.386
0	$\dfrac{4!}{4!0!} = 1$	0

Total $W = 16$

Table 2.2 $N = 10$

n	W	$\ln W$
10	1	0.0
9	10	2.303
8	45	3.807
7	120	4.787
6	210	5.347
5	252	5.529
4	210	5.347
3	120	4.787
2	45	3.807
1	10	2.303
0	1	0.0

Total $W = 1024$

No of head, n in flipping the coin N times

- W (n,N) = N!/N!(N-n)!
- $W(50,100)=1.01 \times 10^{29}$
- $W(25,100) = 2.43 \times 10^{23}$

(a) $N = 10$

(b) $N = 100$

(c) $N = 1000$

Figure 2.6 The multiplicity function W of the number of heads n narrows as the total number of trials N increases.
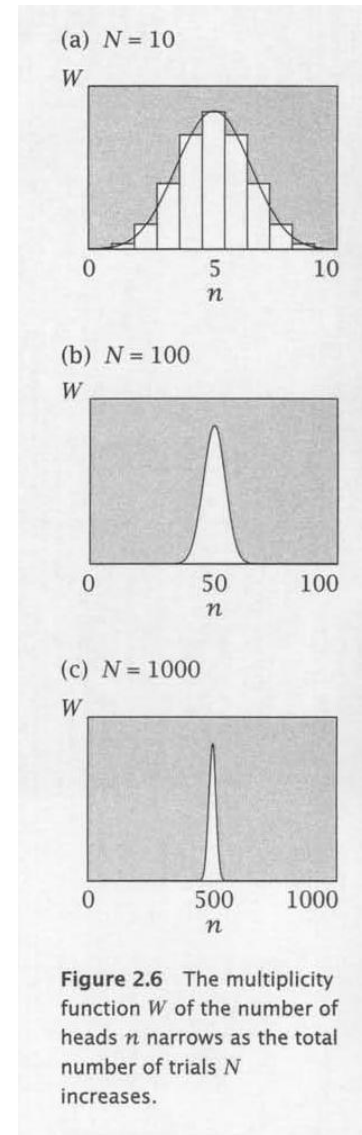
Lattice model to explain why gas molecules spread out into a large volume.

Figure 2.7 For Example 2.2, Case A, three particles are distributed throughout $M_A = 5$ sites. In Case B, three particles are in $M_B = 4$ sites. In Case C the three particles fill $M_C = 3$ sites.

Case	Configuration	Volume
A		5
B		4
C		3

- Multiplicity = W(N,M)
- N= number of particle
- M= number of lattice site
- WA(3,5)=10 possible arrangments
- WB(3,4)=4
- WC(3,3)=1
- If the system has only 3 possible volume, then the probabily is $p_c=1/(10+4+1)$
- $p_B=4/15$ and $p_A=10/15$ (which is most probable.)

Why do materials diffuse?

Figure 2.8 For Example 2.3, Case A, the composition is two ● and two ◐ particles on the left, and two ● and two ◐ on the right. In Case B, the composition is three ● and one ◐ on the left and one ● and three ◐ on the right. In Case C the composition is four ● and zero ◐ on the left, and zero ● and four ◐ on the right.

Why do materials diffuse?

Figure 2.8 For Example 2.3, Case A, the composition is two ● and two ◗ particles on the left, and two ● and two ◗ on the right. In Case B, the composition is three ● and one ◗ on the left and one ● and three ◗ on the right. In Case C the composition is four ● and zero ◗ on the left, and zero ● and four ◗ on the right.

Once again, the statistical mechanical approach is to assume that each spatial configuration (sequence) is equally probable. Find the most probable mixture by maximizing the multiplicity of arrangements. For each given value of left and right compositions, the total multiplicity is the product of the multiplicities for the left and the right sides:

Case A: $W = W(\text{left}) \cdot W(\text{right}) = \dfrac{4!}{2!2!}\dfrac{4!}{2!2!} = 36.$

Case B: $W = W(\text{left}) \cdot W(\text{right}) = \dfrac{4!}{1!3!}\dfrac{4!}{3!1!} = 16.$

Case C: $W = W(\text{left}) \cdot W(\text{right}) = \dfrac{4!}{0!4!}\dfrac{4!}{4!0!} = 1.$

Why do material absorb energy?

- For instance 3 distinguishable particles.

- What is the total number W(U) of ways that the system can partition its energy?

- There are more arrangements of a system that have a high energy than a low energy.
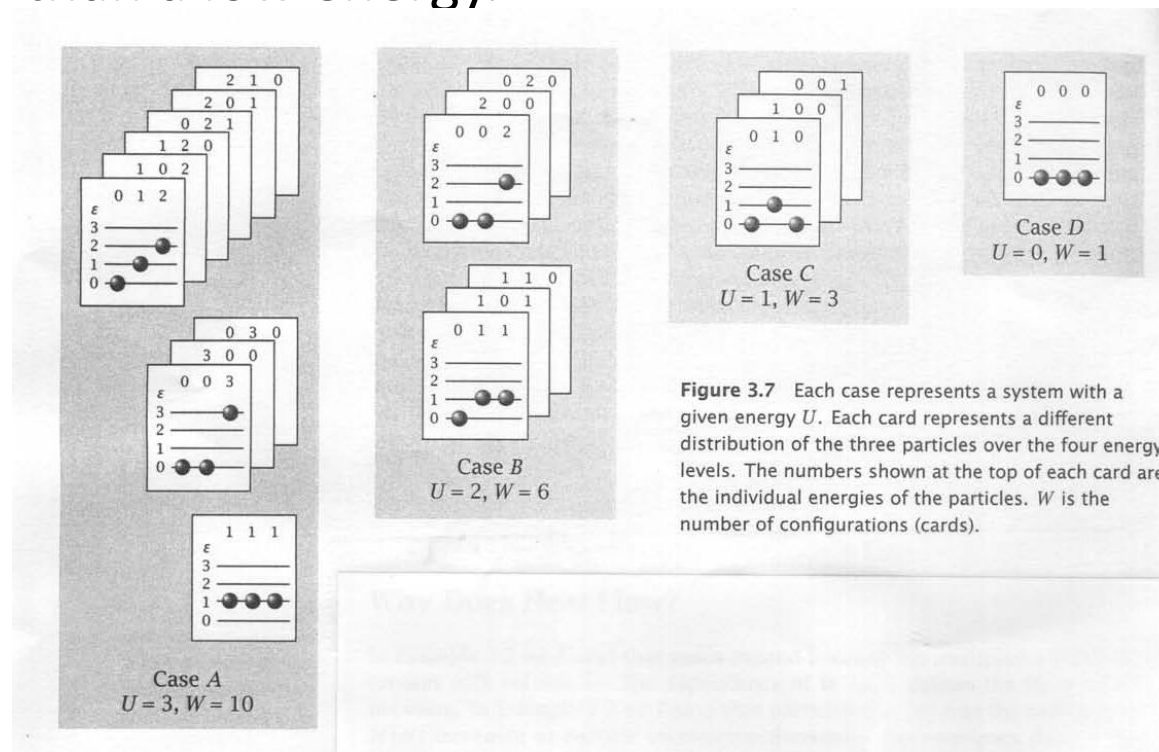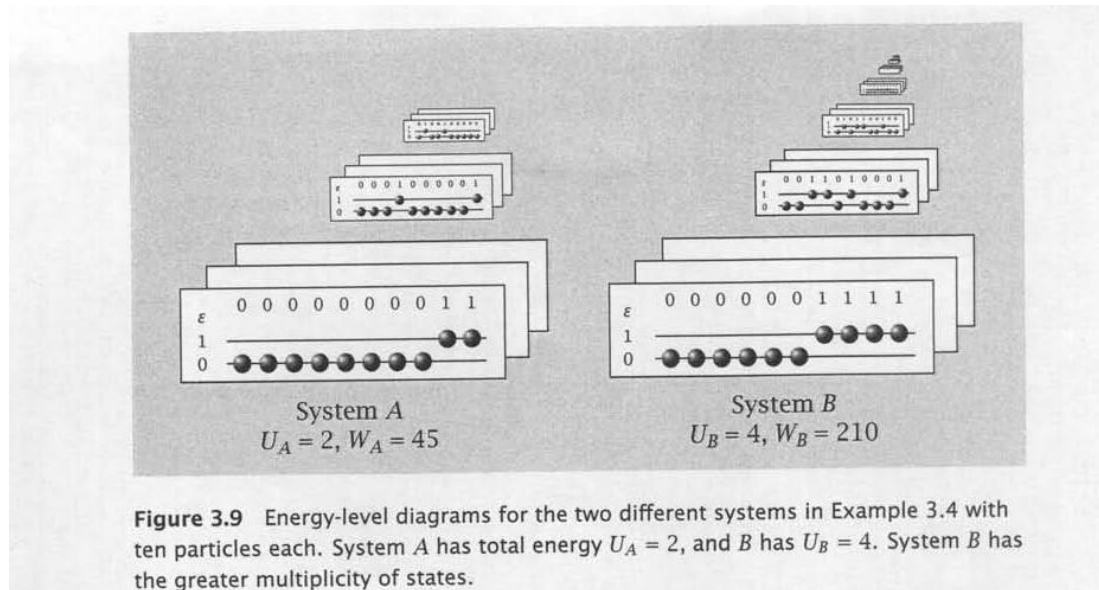

Figure 3.7 Each case represents a system with a given energy U. Each card represents a different distribution of the three particles over the four energy levels. The numbers shown at the top of each card are the individual energies of the particles. W is the number of configurations (cards).

Case A
$U = 3, W = 10$

Case B
$U = 2, W = 6$

Case C
$U = 1, W = 3$

Case D
$U = 0, W = 1$

Why does energy exchange?

Figure 3.9 Energy-level diagrams for the two different systems in Example 3.4 with ten particles each. System A has total energy $U_A = 2$, and B has $U_B = 4$. System B has the greater multiplicity of states.

EXAMPLE 3.4 Why does energy exchange? Consider the two systems, A and B, shown in Figure 3.9. Each system has ten particles and only two energy levels, $\varepsilon = 0$ or $\varepsilon = 1$ for each particle. The binomial statistics of coin flips applies to this simple model.

Systems A and B are identical except that A has less energy than B. System A has two particles in the 'excited state' ($\varepsilon = 1$), and eight in the 'ground state' ($\varepsilon = 0$) so the total energy is $U_A = 2$. System B has four in the excited state and six in the ground state so $U_B = 4$. In this case, heat flows from B (higher energy) to A (lower energy) because doing so increases the multiplicity. Let's see how.

The multiplicities of isolated systems A and B are

$$W_A = \frac{10!}{2!8!} = 45, \quad \text{and} \quad W_B = \frac{10!}{4!6!} = 210.$$

24

Why does energy exchange?

Figure 3.9 Energy-level diagrams for the two different systems in Example 3.4 with ten particles each. System A has total energy $U_A = 2$, and B has $U_B = 4$. System B has the greater multiplicity of states.
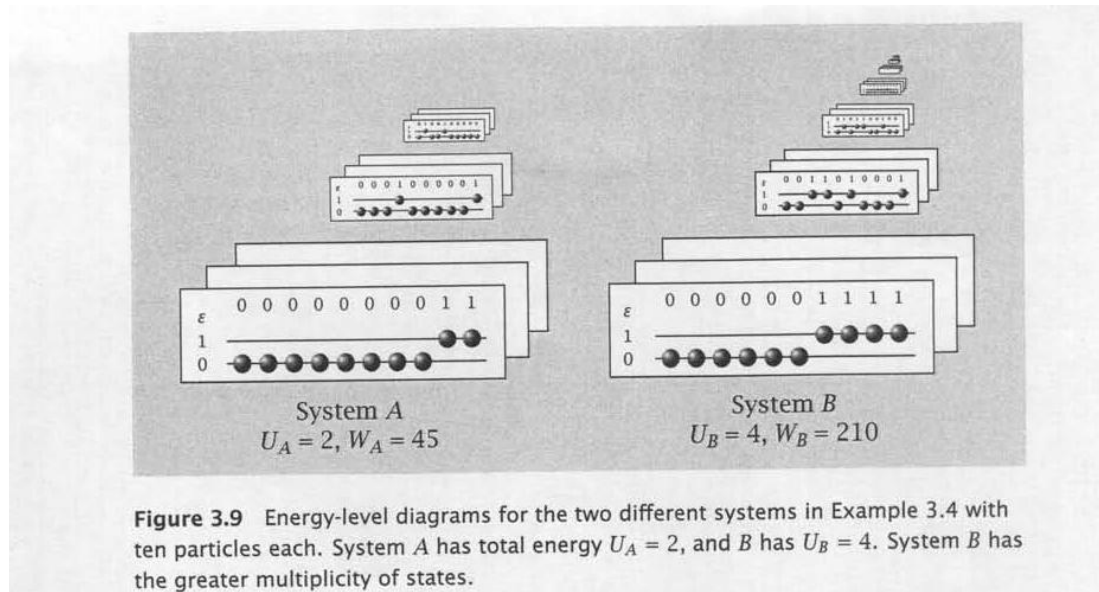
EXAMPLE 3.4 Why does energy exchange? Consider the two systems, A and B, shown in Figure 3.9. Each system has ten particles and only two energy levels, $\varepsilon = 0$ or $\varepsilon = 1$ for each particle. The binomial statistics of coin flips applies to this simple model.

Systems A and B are identical except that A has less energy than B. System A has two particles in the 'excited state' ($\varepsilon = 1$), and eight in the 'ground state' ($\varepsilon = 0$) so the total energy is $U_A = 2$. System B has four in the excited state and six in the ground state so $U_B = 4$. In this case, heat flows from B (higher energy) to A (lower energy) because doing so increases the multiplicity. Let's see how.

The multiplicities of isolated systems A and B are

$$W_A = \frac{10!}{2!8!} = 45, \quad \text{and} \quad W_B = \frac{10!}{4!6!} = 210.$$

If A and B do not exchange energies, the total multiplicity is $W_{total} = W_A W_B = 9450$. Now suppose that you bring A and B into 'thermal contact' so that they can exchange energy. Now the system can change values of U_A and U_B subject to conservation of energy ($U_A + U_B$ will be unchanged). One possibility is $U_A = 3$ and $U_B = 3$. Then the total multiplicity W_{total} will be

$$W_{total} = \frac{10!}{3!7!}\frac{10!}{3!7!} = 14400.$$

This shows that a principle of maximum multiplicity predicts that heat will flow to equalize energies in this case. Consider the alternative. Suppose A were to lower its energy to $U_A = 1$ while B wound up with $U_B = 5$. Then the multiplicity of states would be

$$W_{total} = \frac{10!}{1!9!}\frac{10!}{5!5!} = 2520.$$

A maximum-multiplicity principle predicts that this inequitable distribution is unlikely. That is, heat will not flow from the cold to the hot object.

However, energy does not always flow downhill.

EXAMPLE 3.5 However, energy doesn't always flow downhill. Example 3.4 predicts that energies tend to equalize. But here's a more interesting case that shows that the tendency to maximize multiplicity does not always result in a draining of energy from higher to lower. Again system A has ten particles and an energy $U_A = 2$. However, now system B is smaller, with only four particles, and has energy $U_B = 2$. The energies of A and B are equal. The multiplicity is

$$W = W_A W_B = \frac{10!}{2!8!} \frac{4!}{2!2!} = 45 \times 6 = 270.$$

Now suppose that A and B come into thermal contact and the larger system absorbs energy from the smaller one, so $U_A = 3$ and $U_B = 1$. This causes the multiplicity to increase:

$$W = W_A W_B = \frac{10!}{3!7!} \frac{4!}{1!3!} = 120 \times 4 = 480.$$
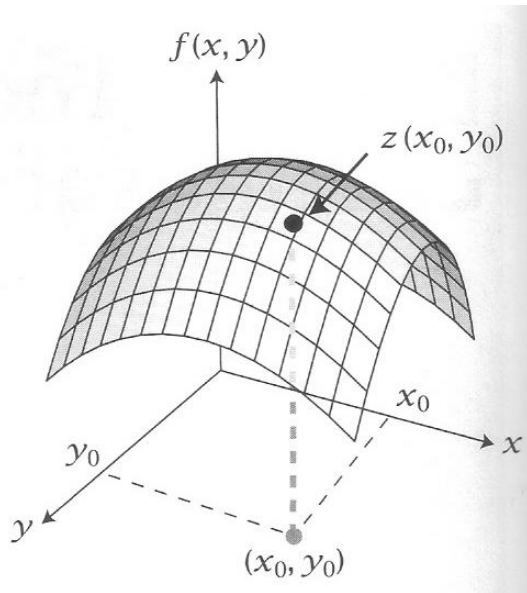
Summary

- Systems tend towards their states of maximum multiplicity (or maximum number of microstates).

- Every microstate is equally probable.

Announcements

- Homework and lecture notes can be downloaded from the website of Ms Ma.

- 1st Homework will be uploaded to the website this Thursday.

- Please work independently.

Math tools to find the extremum of a a multivariate functions

$$\left(\frac{\partial f}{\partial x}\right)_y = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}, \quad \text{and}$$

$$\left(\frac{\partial f}{\partial y}\right)_x = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y},$$

Find the minimum of a paraboloid subject to a constraint.

Figure 5.8 The global minimum of $f(x, y) = x^2 + y^2$ is $(x^*, y^*) = (0, 0)$. But the minimum of $f(x, y)$ on the plane $g(x, y) = x + y = 6$ is $(x^*, y^*) = (3, 3)$.
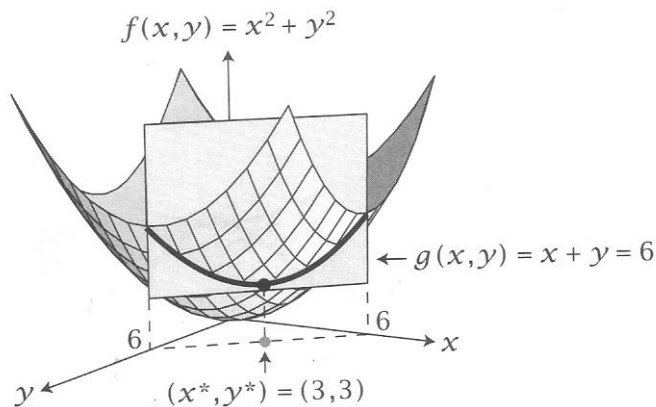
Method of Lagrange multiplier

$$\left(\frac{\partial f}{\partial x}\right)_y = \lambda \left(\frac{\partial g}{\partial x}\right)_y \quad \text{and} \quad \left(\frac{\partial f}{\partial y}\right)_x = \lambda \left(\frac{\partial g}{\partial y}\right)_x.$$

$f(x, y) = x^2 + y^2$

$\leftarrow g(x, y) = x + y = 6$

6

6

x

y

$(x^*, y^*) = (3, 3)$

Method of Lagrange multiplier with 2 constraints.

$$d(f - \lambda g - \beta h) = \sum_{i=1}^{t} \left[\left(\frac{\partial f}{\partial x_i}\right) - \lambda \left(\frac{\partial g}{\partial x_i}\right) - \beta \left(\frac{\partial h}{\partial x_i}\right) \right] dx_i = 0.$$

Classwork

Figure 5.8 The global minimum of $f(x,y) = x^2 + y^2$ is $(x^*, y^*) = (0,0)$. But the minimum of $f(x,y)$ on the plane $g(x,y) = x + y = 6$ is $(x^*, y^*) = (3,3)$.

EXAMPLE 5.6 Finding the minimum of a paraboloid with Lagrange multipliers. Again let's find the minimum of the paraboloid $f(x,y) = x^2 + y^2$ subject to $g(x,y) = x + y = 6$. But now let's use the Lagrange multiplier method. We have $(\partial g/\partial x) = (\partial g/\partial y) = 1$. According to the Lagrange method, the solution will be given by

$$\left(\frac{\partial f}{\partial x}\right)_y = \lambda \left(\frac{\partial g}{\partial x}\right)_y \implies 2x^* = \lambda, \tag{5.23}$$

$$\left(\frac{\partial f}{\partial y}\right)_x = \lambda \left(\frac{\partial g}{\partial y}\right)_x \implies 2y^* = \lambda. \tag{5.24}$$

Combining Equations (5.23) and (5.24), and using $x + y = 6$, gives

$$2x^* = 2y^* \implies y^* = x^* = 3. \tag{5.25}$$

Instantaneous configurations

If, as a result of collisions, the system were to fluctuate between the configurations $\{N,0,0,\ldots\}$ and $\{N-2,2,0,\ldots\}$, it would almost always be found in the second, more likely state (especially if N were large). In other words, a system free to switch between the two configurations would show properties characteristic almost exclusively of the second configuration. A general configuration $\{n_0,n_1,\ldots\}$ can be achieved in W different ways, where W is called the **weight** of the configuration. The weight of the configuration $\{n_0,n_1,\ldots\}$ is given by the expression

$$W = \frac{N!}{n_0! n_1! n_2! \ldots} \tag{16.1}$$

It will turn out to be more convenient to deal with the natural logarithm of the weight, $\ln W$, rather than with the weight itself. We shall therefore need the expression

$$\ln W = \ln \frac{N!}{n_0! n_1! n_2! \ldots} = \ln N! - \ln(n_0! n_1! n_2! \cdots)$$

$$= \ln N! - (\ln n_0! + \ln n_1! + \ln n_2! + \cdots)$$

$$= \ln N! - \sum_i \ln n_i!$$

where in the first line we have used $\ln(x/y) = \ln x - \ln y$ and in the second $\ln xy = \ln x + \ln y$. One reason for introducing $\ln W$ is that it is easier to make approximations. In particular, we can simplify the factorials by using **Stirling's approximation** in the form

$$\ln x! \approx x \ln x - x \tag{16.2}$$

Then the approximate expression for the weight is

$$\ln W = (N \ln N - N) - \sum_i (n_i \ln n_i - n_i) = N \ln N - \sum_i n_i \ln n_i \tag{16.3}$$

Boltzmann distribution

We remarked in Section 16.1 that ln W is easier to handle than W. Therefore, to find the form of the Boltzmann distribution, we look for the condition for ln W being a maximum rather than dealing directly with W. Because ln W depends on all the n_i, when a configuration changes and the n_i change to $n_i + \mathrm{d}n_i$, the function ln W changes to ln $W + \mathrm{d}$ ln W, where

$$\mathrm{d}\ln W = \sum_i \left(\frac{\partial \ln W}{\partial n_i}\right) \mathrm{d}n_i$$

All this expression states is that a change in ln W is the sum of contributions arising from changes in each value of n_i. At a maximum, d ln $W = 0$. However, when the n_i change, they do so subject to the two constraints

$$\sum_i \varepsilon_i \mathrm{d}n_i = 0 \qquad \sum_i \mathrm{d}n_i = 0 \qquad\qquad (16.48)$$

We employ the technique as follows. The two constraints in eqn 16.48 are multiplied by the constants $-\beta$ and α, respectively (the minus sign in $-\beta$ has been included for future convenience), and then added to the expression for d ln W:

$$\mathrm{d}\ln W = \sum_i \left(\frac{\partial \ln W}{\partial n_i}\right) \mathrm{d}n_i + \alpha \sum_i \mathrm{d}n_i - \beta \sum_i \varepsilon_i \mathrm{d}n_i$$

$$= \sum_i \left\{ \left(\frac{\partial \ln W}{\partial n_i}\right) + \alpha - \beta\varepsilon_i \right\} \mathrm{d}n_i$$

Boltzmann distribution

All the dn_i are now treated as independent. Hence the only way of satisfying $d \ln W = 0$ is to require that, for each i,

$$\frac{\partial \ln W}{\partial n_i} + \alpha - \beta \varepsilon_i = 0 \qquad (16.49)$$

when the n_i have their most probable values.

Differentiation of $\ln W$ as given in eqn 16.3 with respect to n_i gives

$$\frac{\partial \ln W}{\partial n_i} = \frac{\partial (N \ln N)}{\partial n_i} - \sum_j \frac{\partial (n_j \ln n_j)}{\partial n_i}$$

The derivative of the first term is obtained as follows:

$$\frac{\partial (N \ln N)}{\partial n_i} = \left(\frac{\partial N}{\partial n_i} \right) \ln N + N \left(\frac{\partial \ln N}{\partial n_i} \right)$$

$$= \ln N + \frac{\partial N}{\partial n_i} = \ln N + 1$$

The $\ln N$ in the first term on the right in the second line arises because $N = n_1 + n_2 + \cdots$ and so the derivative of N with respect to any of the n_i is 1: that is, $\partial N/\partial n_i = 1$. The second term on the right in the second line arises because $\partial(\ln N)/\partial n_i = (1/N)\partial N/\partial n_i$. The final 1 is then obtained in the same way as in the preceding remark, by using $\partial N/\partial n_i = 1$.

For the derivative of the second term we first note that

$$\frac{\partial \ln n_j}{\partial n_i} = \frac{1}{n_j} \left(\frac{\partial n_j}{\partial n_i} \right)$$

Morever, if $i \neq j$, n_j is independent of n_i, so $\partial n_j/\partial n_i = 0$. However, if $i = j$,

$$\frac{\partial n_j}{\partial n_i} = \frac{\partial n_j}{\partial n_j} = 1$$

Therefore,

$$\frac{\partial n_j}{\partial n_i} = \delta_{ij}$$

with δ_{ij} the Kronecker delta ($\delta_{ij} = 1$ if $i = j$, $\delta_{ij} = 0$ otherwise). Then

$$\sum_j \frac{\partial (n_j \ln n_j)}{\partial n_i} = \sum_j \left\{ \left(\frac{\partial n_j}{\partial n_i} \right) \ln n_j + n_j \left(\frac{\partial \ln n_j}{\partial n_i} \right) \right\}$$

$$= \sum_j \left\{ \left(\frac{\partial n_j}{\partial n_i} \right) \ln n_j + \left(\frac{\partial n_j}{\partial n_i} \right) \right\}$$

$$= \sum_j \left(\frac{\partial n_j}{\partial n_i} \right) (\ln n_j + 1)$$

$$= \sum_j \delta_{ij} (\ln n_j + 1) = \ln n_i + 1$$

Boltzmann distribution

and therefore

$$\frac{\partial \ln W}{\partial n_i} = -(\ln n_i + 1) + (\ln N + 1) = -\ln \frac{n_i}{N}$$

It follows from eqn 16.49 that

$$-\ln \frac{n_i}{N} + \alpha - \beta \varepsilon_i = 0$$

and therefore that

$$\frac{n_i}{N} = e^{\alpha - \beta \varepsilon_i}$$

At this stage we note that

$$N = \sum_i n_i = \sum_i N e^{\alpha - \beta \varepsilon_i} = N e^{\alpha} \sum_i e^{-\beta \varepsilon_i}$$

Because the N cancels on each side of this equality, it follows that

$$e^{\alpha} = \frac{1}{\sum_j e^{-\beta \varepsilon_j}} \qquad\qquad (16.50)$$

and

$$\frac{n_i}{N} = e^{\alpha - \beta \varepsilon_i} = e^{\alpha} e^{-\beta \varepsilon_i} = \frac{1}{\sum_j e^{-\beta \varepsilon_j}} e^{-\beta \varepsilon_i}$$

Pressure of atmosphere

EXAMPLE 10.1 Barometric pressure of the atmosphere. The energy ε of a gas molecule in the earth's gravitational field is a function of altitude z:

$$\varepsilon(z) = mgz, \tag{10.12}$$

where g is the gravitational constant and m is the molecular mass. In this case, the energy is a continuous function (of z), not a discrete ladder, but Boltzmann's law still applies. We assume that the atmosphere is in equilibrium (valid only approximately) and is at constant temperature. The number of molecules $N(z)$ at altitude z relative to the number $N(0)$ at sea level is given by the Boltzmann law,

$$\frac{N(z)}{N(0)} = e^{-\varepsilon(z)/kT} = e^{-mgz/kT}. \tag{10.13}$$

If the gas is ideal and the temperature is constant, the pressure $p(z)$ is proportional to the number of molecules per unit volume, so the pressure should decrease exponentially with altitude:

$$\frac{p(z)}{p(0)} = \frac{N(z)kT/V}{N(0)kT/V} = \frac{N(z)}{N(0)} = e^{-mgz/kT}. \tag{10.14}$$

Figure 10.3 shows experimental evidence that the temperature is reasonably constant for the earth's atmosphere up to about $100\,\text{km}$ above the earth's surface, and it shows that the pressure decreases exponentially with altitude, as predicted. Above about $100\,\text{km}$, the equilibrium assumption no longer holds because the atmosphere becomes too thin for normal wind turbulence to mix the gases, and the temperature is no longer independent of altitude.
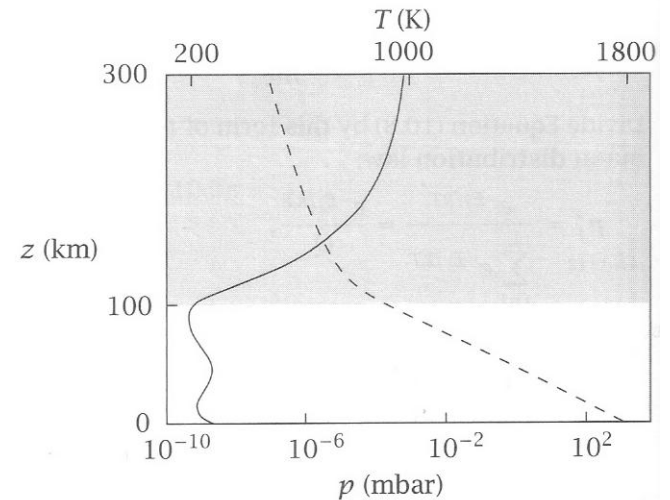

Figure 10.3 The temperature T (———) of the atmosphere is approximately constant up to about $100\,\text{km}$ in altitude z. The pressure p (– – –) decreases exponentially with altitude, following the Boltzmann law. Source: ML Salby, *Fundamentals of Atmospheric Physics*, Academic Press, San Diego, 1996. Data are from: *US Standard Atmosphere*, NOAA, US Air Force, US Government Printing Office, NOAA-S/T 76-1562, Washington, DC, 1976.

Partition function of a uniform ladder of energy levels

Example 16.2 *Evaluating the partition function for a uniform ladder of energy levels*

Evaluate the partition function for a molecule with an infinite number of equally spaced nondegenerate energy levels (Fig. 16.3). These levels can be thought of as the vibrational energy levels of a diatomic molecule in the harmonic approximation.

Method We expect the partition function to increase from 1 at $T = 0$ and approach infinity as T to ∞. To evaluate eqn 16.8 explicitly, note that

$$1 + x + x^2 + \cdots = \frac{1}{1-x}$$

Answer If the separation of neighbouring levels is ε, the partition function is

$$q = 1 + e^{-\beta\varepsilon} + e^{-2\beta\varepsilon} + \cdots = 1 + e^{-\beta\varepsilon} + (e^{-\beta\varepsilon})^2 + \cdots = \frac{1}{1 - e^{-\beta\varepsilon}}$$

This expression is plotted in Fig. 16.4: notice that, as anticipated, q rises from 1 to infinity as the temperature is raised.
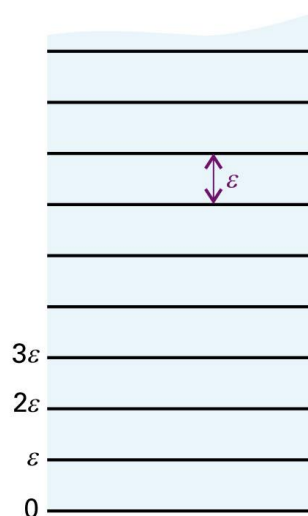
Fig. 16.3 The equally spaced infinite array of energy levels used in the calculation of the partition function. A harmonic oscillator has the same spectrum of levels.

To find a biased die.

$$p_i^* = \frac{p_i^*}{\sum\limits_{i=1}^{t} p_i^*} = \frac{e^{(-1-\alpha)}e^{-\beta \varepsilon_i}}{\sum\limits_{i=1}^{t} e^{(-1-\alpha)}e^{-\beta \varepsilon_i}} = \frac{e^{-\beta \varepsilon_i}}{\sum\limits_{i=1}^{t} e^{-\beta \varepsilon_i}}. \qquad (6.16)$$

In statistical mechanics, this is called the **Boltzmann distribution law** and the quantity in the denominator is called the **partition function** q.

$$q = \sum_{i=1}^{t} e^{-\beta \varepsilon_i}. \qquad (6.17)$$

Using Equations (6.11) and (6.16) you can express the average score per roll $\langle \varepsilon \rangle$ (Equation (6.9)) in terms of the distribution,

$$\langle \varepsilon \rangle = \sum_{i=1}^{t} \varepsilon_i p_i^* = \frac{1}{q} \sum_{i}^{t} \varepsilon_i e^{-\beta \varepsilon_i}. \qquad (6.18)$$

The next two examples show how Equation (6.18) predicts all t of the p_i^*'s from the one known quantity, the average score.

EXAMPLE 6.3 Finding bias in dice by using the exponential distribution law. Here we illustrate how to predict the maximum entropy distribution when an average score is known. Suppose a die has $t = 6$ faces and the scores equal the face indices, $\varepsilon(i) = i$. Let $x = e^{-\beta}$. Then Equation (6.17) gives $q = x + x^2 + x^3 + x^4 + x^5 + x^6$, and Equation (6.16) gives

$$p_i^* = \frac{x^i}{\sum\limits_{i=1}^{6} x^i} = \frac{x^i}{x + x^2 + x^3 + x^4 + x^5 + x^6}. \qquad (6.19)$$

From the constraint Equation (6.18), you have

$$\langle \varepsilon \rangle = \sum_{i=1}^{6} i p_i^* = \frac{x + 2x^2 + 3x^3 + 4x^4 + 5x^5 + 6x^6}{x + x^2 + x^3 + x^4 + x^5 + x^6}, \qquad (6.20)$$

You have a polynomial, Equation (6.20), that you must solve for the one unknown x (a method for solving polynomials like Equation (6.20) is given on page 551. You begin with knowledge of $\langle \varepsilon \rangle$. Compute the value x^* that solves Equation (6.20). Then substitute x^* into Equations (6.19) to give the distribution function $(p_1^*, p_2^*, \ldots, p_t^*)$.

For example, if you observe the average score $\langle \varepsilon \rangle = 3.5$, then $x = 1$ satisfies Equation (6.20), predicting $p_i^* = 1/6$ for all i, indicating that the die is unbiased and has a flat distribution (see Figure 6.3(a)).

Instead, you observe the average score is $\langle \varepsilon \rangle = 3.0$, then $x = 0.84$ satisfies Equation (6.20), and you have $q = 0.84 + 0.84^2 + 0.84^3 + 0.84^4 + 0.84^5 + 0.84^6 = 3.41$. The probabilities are $p_1 = 0.84/3.41 = 0.25$, $p_2 = 0.84^2/3.41 = 0.21$, $p_3 = 0.84^3/3.41 = 0.17$, and so on, as shown in Figure 6.3(b).
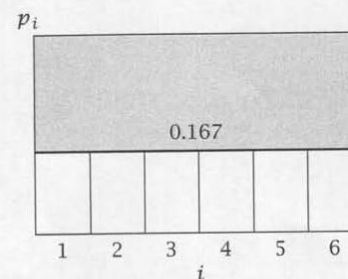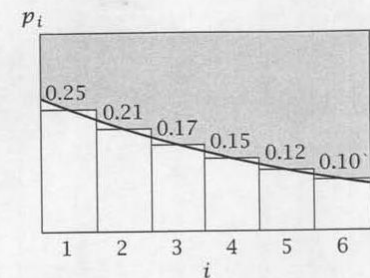

Figure 6.3 The probabilities of dice outcomes for known average scores. (a) If the average score per roll is $\langle \varepsilon \rangle = 3.5$, then $x = 1$ and all outcomes are equally probable, predicting that the die is unbiased. (b) If the average score is low ($\langle \varepsilon \rangle = 3.0$, $x = 0.84$), maximum entropy predicts an exponentially diminishing distribution. (c) If the average score is high ($\langle \varepsilon \rangle = 4.0$, $x = 1.19$), maximum entropy implies an exponentially increasing distribution.