Discussion 8      Mar 12, 2019

We will discuss generalized canonical formalisms and information.

1. **[Pressure ensemble]**
Starting from the microcanonical approach, we go to the canonical formalism that allows exchange of energy between the system and the environment (thermostat).

Compare the following formulas:

\[
S = k_B \log W(E,V), \quad \text{(DH8.1)}
\]

\[
-A/T = \max_E (S - E) = k_B \log Z(T,V). \quad \text{(DH8.2)}
\]

The first equality in (DH8.2) is due to the standard Legendre transformation applied to \( E \): 

\[
-A = \max_S [ST - E] = \max_E [ST - E] \quad (T \text{ is a constant}).
\]

Boltzmann’s principle (DH8.1) means that the first equality (which is a thermodynamic relation) in (DH8.2) is equivalent to (recall \( 1/k_B T = \beta \))

\[
e^{-\beta A} = \max_E e^{S/k_B - \beta E} = \max_E [W(E,V)e^{-E/k_B T}]. \quad \text{(DH8.3)}
\]

The key to the ‘general’ ensemble method is the ‘max-sum’ correspondence:

\[
\max_E [W(E,V)e^{-E/k_B T}] = \sum_E W(E,V)e^{-E/k_B T} = Z(T,V) \quad \text{(DH8.4)}
\]

with a relative error of order \( \log N/N \).

---

1. If you wish to use the Legendre transformation applied to entropy (or \(-S\), which is convex), we first look at the Gibbs relation

\[
d(-S) = \left(-\frac{1}{T}\right)dE + \left(-\frac{P}{T}\right)dV.
\]

so a Legendre transformation applied to the convex function \(-S\) can be written as

\[
\max_E \left[\left(-\frac{1}{T}\right)dE - (-S)\right] = \max_E \{S - E/T\} = -A/T,
\]

which is (DH8.2).

2. The \( E \) integral may be understood as summation over energy shells.

3. As you see, for statistical mechanics \( S, -A/T, \) etc. (the so-called Massieu functions), are much
Let us consider an ideal gas under constant temperature and pressure. What is the most convenient thermodynamic potential?

Write down the partition function $Y$ that directly gives the thermodynamic potential in (1) as $-k_B T \log Y$.

Compute $Y$ for a monatomic ideal gas. You may use its canonical partition function

$$Z(T,V) = \frac{1}{N!} \left( \frac{2\pi mk_B T}{\hbar^2} \right)^{3N/2} V^N.$$ (DH8.6)

Recover the ideal gas law.

**Soln.**

We know the answer is $G = E - ST + PV$, the Gibbs free energy. If you wish to retrace all the argument to derive the ‘generalized’ canonical formalism for the present case, read the following.

To maintain the temperature and pressure of the system, heat and volume work should be freely exchanged with its environment. Therefore, we pay attention to the energy change

$$\Delta E - \Delta Q + \Delta(PV) = \Delta E - T\Delta S + \Delta(PV) = \Delta G,$$ (DH8.7)

where $G$ is the Gibbs free energy as we already know. The Legendre transformation we use is

$$-G = \max_{S,V} [ST + (-P)V - E]$$ (DH8.8)

or

$$-\frac{G}{T} = \max_{E,V} \left[ S - \frac{P}{T} V - \frac{E}{T} \right].$$ (DH8.9)

This is directly related to the Legendre transformation of entropy (recall that $-S$ is convex; see Footnote 1):

$$-\frac{G}{T} = \max_{E,V} \left[ -\frac{P}{T} V - \frac{E}{T} - (-S) \right],$$ (DH8.10)

more natural than the energies (thermodynamic potentials). Also look at the Gibbs relations to see what really the independent variables are for statistical mechanics:

$$d \left( S - \frac{E}{T} \right) = -d \frac{A}{T} = -Ed T + \frac{P}{T} dV + \cdots.$$ (DH8.5)

Note that this immediately gives us the Gibbs-Helmholtz formula.
because, for each $V$, $S$ and $E$ are one-to-one correspondent.

(2) Therefore, we obtain
\[
e^{-\beta G} = \max_{E,V} e^{S/k_B - \beta(E + PV)},
\]
(DH8.11)

Using Boltzmann’s principle, we get
\[
e^{-\beta G} = \max_{E,V} \left[ W(E, V)e^{-\beta(E + PV)} \right],
\]
(DH8.12)

where $W(E, V)$ is the total ‘number’ of microstates compatible with the thermodynamic state $(E, V)$; $W$ is often called the microcanonical partition function. We can also write the above formula as
\[
e^{-\beta G} = \max_{V} \left[ Z(T, V)e^{-\beta PV} \right],
\]
(DH8.13)

because (DH8.8) implies
\[
-G = \max_{V} \left[ (-P)V + \max_{S}[S(T - E)] = \max_{V} [(-P)V - A]\right]
\]
and because we know
\[
e^{-\beta A} = \max_{E} \left[ W(E, V)e^{-\beta E} \right] = Z(T, V).
\]
(DH8.14)

Needless to say, we always ignore the relative errors of $O[\log N/N]$, which is very small already for $N \sim 1000$.

(DH8.13) implies (‘max-sum’ correspondence, although the sum here is written as an integral)
\[
e^{-\beta G} = \int dV Z(T, V, X)e^{-\beta PV} = Y(T, P),
\]
(DH8.15)
or you may use, shortcutting the intermediate ‘max-sum’ correspondence,
\[
e^{-\beta G} = e^{-\beta(A + PV)} \Rightarrow e^{-\beta G} = \sum_{V} Z(T, V)e^{-\beta PV} = Y(T, P)
\]
(DH8.16)
to memorize the practical rule.

(3) The canonical partition function is given as
\[
Z(T, V) = \frac{1}{N!} \left[ \left( \frac{2\pi m k_B T}{\hbar^2} \right)^{3/2} \right]^N V^N,
\]
(DH8.17)

\footnote{Notice that $A$ is a convex function of $V$, because $E$ is. $A$ is a concave function of $T$. Thus, $A$ is a rather complicated multivariate function.}
so we need the following integral:

\[
\int dV \frac{1}{N!} V^N e^{-\beta PV} = \frac{1}{N!(\beta P)^{N+1}} \int dx x^N e^{-x} = \frac{1}{N!(\beta P)^{N+1}} \Gamma(N+1) = \left(\frac{k_B T}{P}\right)^N.
\]

Therefore, (I ignore the difference between \(N\) and \(N+1\))

\[
Y(T, P) = \left[ \left(\frac{2\pi m}{\hbar^2}\right)^{3/2} \frac{(k_B T)^{5/2}}{P} \right]^N.
\]

(4) Thus,

\[
G = -k_B T \log Y = Nk_B T \log P + \cdots,
\]

We know \(dG = V dP - SdT + \cdots\), so

\[
V = \left(\frac{\partial G}{\partial P}\right)_T = \frac{Nk_B T}{P}.
\]

Incidentally, the relation between \(Y\) and \(Z\) is, as seen from (DH8.13),

\[
PV = k_B T \log \frac{Z(T, V)}{Y(T, P)} = k_B T \log \frac{V^N/N!}{(k_B T/P)^N} = Nk_B T \log \frac{PV e^{Nk_B T}}{Nk_B T}.
\]

Here we have used \(N! \approx (N/e)^N\). This implies that \(x = PV/Nk_B T\) satisfies

\[
x = 1 + \log x.
\]

\(x = 1\) is the unique real solution (as you can easily see graphically).

2 [Elementary problem about spin system]

“Due to the ligand field the degeneracy of the \(d\)-orbitals of the chromium ion \(\text{Cr}^{3+}\) is lifted, and the spin Hamiltonian has the following form

\[
H = D(S^2 - S(S+1)/2),
\]

where \(D > 0\) is a constant with \(S = 3/2\) (the cation is in the term \(^4\text{F}_{3/2}\)).”

This is the way a spin state question is asked, e.g., in a Qual (the problem is an actual qual problem). However, to answer the statistical-mechanical questions, you need not understand the quantum mechanical setup, but you have only to understand the following facts about the system:
Each ion has states with energies $\epsilon = 3D/8$ and $\epsilon = -13D/8$ and both are doubly degenerate (that is, there are two states each with either of $\epsilon$).

(0) Why can you apply statistical mechanics to this ‘single’ ion?

(1) Compute the occupation probability of each state at temperature $T$ (you may use the standard notation $\beta = 1/k_BT$).

(2) Calculate the entropy.

(3) At high temperatures approximately the specific heat of the system is approximately $C = k_B(T_0/T)^2$ with $T_0 = 0.18$ K. Determine $D/k_B$ in K.

Soln

(0) Statistical mechanics can compute the probability of a set of microstates

(i) that may be specified by a small number of conditions, and

(ii) that contains $N$ microstates such that $\log N$ is extensive ($O[N]$, where $N$ is the total number of particles).

In particular, if all the $N$ microscopic variables are statistically independent as in the systems consisting of non-interacting particles, although a single particle state itself is not macroscopic, a set of microstates in which the particle assumes a particular single particle state obviously satisfies (i) and (ii). The outcome looks as if we can apply statistical mechanics directly to a single molecular state: Let $h_i$ be the single particle Hamiltonian of the $i$th particle and the specified single particle be $i = 0$. Then, the (canonical) probability for a set of microstates in which $h_0 = \epsilon$ is given by

$$P(h_0 = \epsilon) = \frac{1}{Z} \sum_{\{h_i\}_{i \neq 0, h_0 = \epsilon}} e^{-\beta \sum_i h_i} \propto e^{-\beta \epsilon}. \quad \text{(DH8.25)}$$

(1) There are 4 states, but there are only two energy levels with $\epsilon = 3D/8$ and $-13D/8$. Therefore, a one-particle (or single-spin) state with $\epsilon = 3D/8$ is with

$$p = \frac{e^{-3\beta D/8}}{2(e^{-3\beta D/8} + e^{13\beta D/8})} = \frac{1}{2(1 + e^{2\beta D})}. \quad \text{(DH8.26)}$$

Do not forget the $1/2$ factor. The one-particle state with $\epsilon = -13D/8$ is with

$$p' = \frac{e^{13\beta D/8}}{2(e^{-3\beta D/8} + e^{13\beta D/8})} = \frac{e^{2\beta D}}{2(1 + e^{2\beta D})}. \quad \text{(DH8.27)}$$
(2) The easiest method is to use the Shannon formula.\(^5\)

\[
S = -2k_B \left[ \frac{1}{2(1+x)} \log \frac{1}{2(1+x)} + \frac{x}{2(1+x)} \log \frac{x}{2(1+x)} \right] = k_B \left\{ \log[2(1+x)] - \frac{x}{1+x} \log x \right\},
\]

(DH8.28)

where \(x = e^{2\beta D}\).

(3) Setting \(x\) as above, we have

\[
C = T \frac{dS}{dT} = -(2D\beta) \frac{dS}{d2D\beta} = -2D\beta \frac{dx}{d2D\beta} \frac{dS}{dx} = -2D\beta x \frac{dS}{dx} = k_B(2D\beta)^2 \frac{x}{(1+x)^2}.
\]

(DH8.29)

That is, for large \(T\) (small \(\beta\)), \(x \simeq 1\), so

\[
C = k_B(D/k_B)^2/T^2.
\]

(DH8.30)

Therefore, \(D/k_B = T_0\) or \(D/k_B\) is 0.18 K.

3 [Collection of permanent dipoles]

Let us consider a collection of non-interacting electric dipoles \(\{p_i\}\) sitting on the lattice points \(\{i\}\) as illustrated in Fig. DH8.1. We ignore its rotational kinetic energy, so the system Hamiltonian (the total intrinsic energy) is 0.

\begin{figure} [h]
\centering
\includegraphics[width=0.5\textwidth]{FigDH8_1.png}
\caption{Dipole moments sitting on a lattice (the dipoles are actually three-dimensional)}
\end{figure}

A permanent electric dipole \(p\) has a potential energy \(u = -p \cdot E\), if an external\(^5\)There are several ways to compute entropy. If you know probability explicitly, the Shannon formula may be useful. In this case, you must not forget that the sum is over the elementary events. The microcanonical way is probably the least useful in practice. When you compute \(S\) from the canonical ensemble, use \(S = (E - A)/T\) with \(E\) being calculated by the Gibbs-Helmholtz relation \([\partial(A/T)/\partial(1/T)]_V = -\partial \log Z/\partial \beta = E\). Do NOT, in practice, use \(S = -(\partial A/\partial T)_V\) directly.
electric field $E$ is imposed. Thus, the total potential energy of the system in the electric field reads
\[ U = -E \cdot P, \]  
where $P$ is the electric polarization defined as
\[ P = \sum_i p_i. \]

We wish to obtain $P$ ($\approx$ its expectation value thanks to the LLN) as a function of $T$ and $E$.

(0) The ‘standard’ statistical mechanical textbooks proceed as follows. The system Hamiltonian is
\[ 'H' = -E \cdot P. \]

Therefore, the internal energy is its (e.g., canonical) average:
\[ E = \langle 'H' \rangle. \]

The (canonical) partition function is thus given by
\[ 'Z' = \left[ \int d^3p e^{\beta E \cdot p} \right]^N. \]

From this, we obtain the (Helmholtz) free energy $A$ as usual: \[ 'A' = -k_B T \log 'Z', \]
whose Gibbs relation is understood to be
\[ d' A' = -SdT - P \cdot dE. \]

What is wrong (at least awkward) with the above logic?\(^6\)

Let us proceed according to the correct statistical thermodynamics.
(1) What is the thermodynamic coordinates for the system?
(2) What is its internal energy?
(3) What is the definition of the microcanonical ensemble for the system?

\(^6\)However, most textbooks avoid explicit embarrassment by not mentioning the names of the free energy explicitly and not writing the Gibbs relations explicitly; basically evading thermodynamics. Garrod, Wolfe, Reichl, Greiner et al., Le Bellac et al., etc., follow this line (probably without never thinking). Therefore, in practice, you should look at the definition of the ‘Hamiltonian’ and the partition functions, and then determine which Gibbs relation you should use when you read usual books or attend usual courses. Perhaps, to be successful, don’t be too serious.
(4) Fixing \( P \) is inconvenient. Also we do not wish to put the system in an adiabatic condition, so we wish to keep the system temperature constant (isothermal). Then, for our system what is the convenient thermodynamic potential \( \Phi \)? Also write down the Gibbs relation for \( \Phi \).

(5) What sort of statistical ensemble you should employ to obtain \( \Phi \) directly? That is, what is the corresponding partition function \( \tilde{Z} \)?

(6) Compute \( \Phi = -k_B T \log \tilde{Z} \).

(7) Compute \( P \) as a function of \( E \) and \( T \).

**Soln.**

(0) The key physics observations are:

(i) The internal energy is the energy stored in the system itself,

(ii) The potential energy like \( U \) in this problem (or \( -B \cdot M \) for magnetic systems) is not stored solely in the system itself, but it is stored in the ‘relationship’ between the system and the external field or the device making the external field (just as the volume energy \( PV \); it is there only because something (external to the system) maintains \( P \)).

Thus, ‘\( H \)’ is not the system Hamiltonian.\(^8\) In this case the system Hamiltonian is zero, since we ignore the kinetic energy and since the dipoles are not interacting with each other. Therefore, inevitably \( E = 0 \); since ‘\( H \)’ is not the system Hamiltonian, ‘\( E \)’ cannot be the system internal energy (but is one of the generalized Enthalpies as we will see later).

This misidentification of energy (or the system Hamiltonian) totally screws up the statistical thermodynamic description of the system. Needless to say ‘\( Z \)’ is not the usual canonical partition function, since ‘\( H \)’ is not the system Hamiltonian. If you can declare ‘\( A \)’ to be the system Helmholtz free energy \( A = E - TS \), then the Gibbs relation

\[
d'A = TdS + E \cdot dP
\]

must hold, but in ‘\( A \)’ \( P \) does not appear anywhere (it was averaged or summed away when you compute ‘\( Z \)’). Therefore, in the standard books they ‘declare’ that this ‘Helmholtz free energy’ obeys (DH8.36) to ‘correct’ wrong identifications of quantities.\(^9\)

---

\(^7\)Recall the ‘max-sum’ correspondence.

\(^8\)You could say it is the Hamiltonian of the system and a part of the electric field modified by the presence of the dipoles.

\(^9\)A professor told me that good students should know the difference. Yes, but I believe it is better to point out the conceptual mistake in the usual books and to correct it.
(1) Generally, it is the (internal) energy $E$ and $P$ (recall that they must be energy and work coordinates that are extensive). In the present case $E = 0$ because the system Hamiltonian is zero. Thus, actually, there is only one coordinate $P$ because of the artificial simplification of the problem that has no interactions among different lattice sites.

(2) This is already answered. $E = 0$, invariant. Thus, the Gibbs relation is actually

$$0 = TdS + E \cdot dP,$$

but you may identify this with $dE$ as usual (while noting that it is always 0).

(3) Since the thermodynamic coordinates are $E (= 0)$ and $P$, the microcanonical ensemble consists of microstates with a fixed $P$. Accordingly, the microcanonical partition function $W(E, P)$ (actually, $W(0, P)$) is the total number of microstates with $P$ (with a leeway).

(4) Since we do not wish to fix $P$, we must allow its change. If there is an external electric field $E$, we must allow the system (+ the field) to change its interaction energy freely. Also we wish to allow the system to exchange heat freely with its environment. Thus, instead of the total energy $E + U$ (though not of the system), we should study $\Phi = (E - Q) + U$ (actually a Legendre transform of $E$ as we see in (5)), where $U = -E \cdot P$, the total potential energy. The Gibbs relation for $\Phi$ reads

$$d\Phi = d(A - E \cdot P) = -SdT - P \cdot dE,$$

where $A$ is not the fake Helmholtz free energy (= ‘$A$’) but the true Helmholtz free energy of the system.

(5) The Legendre transformation used in (4) is actually

$$-\Phi = \max_P [E \cdot P + ST - E]$$

energetically, or (recall $d(-S) = (-1/T)dE + (E/T) \cdot dP$ and $-S$ is convex).

$$-\frac{\Phi}{T} = \max_P \left[ \frac{E}{T} \cdot P - \frac{E}{T} - (-S) \right].$$

However, since $E = 0$ definitely in our system, you may drop it from the above formulas.

---

10 More precisely, we must allow some macroscopic leeway such that $P$ is in a (macroscopically) small volume element of $P$. 
From either of the above relations we can write
\[ -\beta \Phi = \max_P \left[ \frac{S}{k_B} + \beta \mathbf{E} \cdot \mathbf{P} \right]. \quad \text{(DH8.42)} \]
or
\[ e^{-\beta \Phi} = \max_P W(0, \mathbf{P}) e^{\beta \mathbf{E} \cdot \mathbf{P}}. \quad \text{(DH8.43)} \]
Thus, with the relative error of \( O[\log N/N] \), we can use the ‘max-sum’ correspondence:
\[ e^{-\beta \Phi} = \sum_P W(0, \mathbf{P}) e^{\beta \mathbf{E} \cdot \mathbf{P}} \equiv \tilde{Z}(T, \mathbf{E}). \quad \text{(DH8.44)} \]
Since \( E \) does not change, you might wonder how we can introduce \( T \) into the description. Although we have \( \beta \), it appears only in combination with \( E \) (and as \( \Phi/T \)). Thus, the entropic form of the Gibbs relation \( dS = -(E/T) \cdot dP \) that cannot define \( T \) alone causes no difficulty. Even though we write \( \tilde{Z}(T, \mathbf{E}) \), it is actually a function of \( E/T \).

(6) Since the summation or the integration over \( \mathbf{P} \) may be decomposed into the sums or integrals for individual molecular dipoles, we can write
\[ \tilde{Z} = \left[ \sum_{\mathbf{p}_1} e^{\beta \mathbf{E} \cdot \mathbf{p}_1} \right]^N. \quad \text{(DH8.45)} \]
Here, the first dipole is used as the representative.

Thus, we have only to compute\(^{11}\)
\[ z(\mathbf{E}) = \int d\mathbf{e} e^{\beta p \mathbf{e} \cdot \mathbf{E}} = \int d\mathbf{e} e^{\beta p \mathbf{E} \cos \theta}, \quad \text{(DH8.46)} \]
where \( \mathbf{e} \) is the directional unit vector of the dipole moment with respect to the electric field direction \( (\mathbf{p} = p \mathbf{e} \) with \( p = |\mathbf{p}|) \), \( E = |\mathbf{E}| \), and the angle between \( \mathbf{E} \) and \( \mathbf{p} \) (or \( \mathbf{e} \)) is \( \theta \).

The integration is on the unit sphere and we can compute the ‘single-body’ or ‘single-site’ canonical partition function as
\[ z(\mathbf{E}) = 2\pi \int_0^\pi d\theta \sin \theta e^{\beta p E \cos \theta} = 2\pi \int_{-1}^1 dx e^{\beta p Ex} = \frac{4\pi}{\beta p E} \sinh \beta p E. \quad \text{(DH8.47)} \]
\(^{11}\)The summation is actually over the directions of the dipole, so the following integral is over the directions, but you need not worry about this and may integrate over \( \mathbf{p} \); \( |\mathbf{p}| \) is constant, so the resultant \( \Phi \) is shifted only by a constant.
Therefore, we have arrived at
\[ \Phi = -Nk_BT \log \left( \frac{4\pi}{\beta pE} \sinh \beta pE \right). \quad (DH8.48) \]

(7) The Gibbs relation \( d\Phi = -SdT - \mathbf{P} \cdot d\mathbf{E} \) (see (4)) tells us
\[ \mathbf{P} = - \left( \frac{\partial \Phi}{\partial \mathbf{E}} \right)_T. \quad (DH8.49) \]

From the structure of \( z \) we can immediately see
\[ \langle p \rangle = k_BT \frac{\partial}{\partial \mathbf{E}} \log z(\mathbf{E}) = p \frac{\partial}{\partial \beta_pE} \log \left( \frac{\sinh \beta_pE}{\beta_pE} \right) = pL(\beta_pE) \frac{\partial E}{\partial \mathbf{E}}, \quad (DH8.50) \]

where \( L(x) \) is the Langevin function
\[ L(x) = \frac{d}{dx} (\log \sinh x - \log x) = \coth x - \frac{1}{x}. \quad (DH8.51) \]

The last derivative with respect to \( E \) (that is a kind of gradient) may be most easily computed from the differentiation of \( E^2 = \mathbf{E} \cdot \mathbf{E} \):
\[ EdE = \mathbf{E} \cdot d\mathbf{E}. \quad (DH8.52) \]

From this we get (or see (DH8.57))
\[ \frac{\partial E}{\partial \mathbf{E}} = \frac{\mathbf{E}}{E}. \quad (DH8.53) \]

Thus,
\[ \mathbf{P} = N\langle \mathbf{p} \rangle = pNL(\beta_pE) \frac{\mathbf{E}}{E}. \quad (DH8.54) \]

If you wish to do all the calculation componentwisely, we can proceed as follows.
\[ P_i = \left( \frac{\partial \Phi}{\partial E_i} \right)_T = Nk_BT \frac{\partial}{\partial E_i} \log \frac{\sinh \beta_pE}{\beta_pE} = pN \frac{\partial}{\partial \beta_pE_i} \log \frac{\sinh \beta_pE}{\beta_pE} \]
\[ = pNL(\beta_pE) \frac{\partial E_i}{\partial \mathbf{E}} = pNL(\beta_pE) \frac{E_i}{E}. \quad (DH8.55) \]
This is just (DH8.54). Here, we used
\[
\frac{\partial E}{\partial E_i} = \frac{\partial |E|}{\partial E_i} = \frac{\partial \sqrt{\sum_i E_i^2}}{\partial E_i} = \frac{2E_i}{2\sqrt{\sum_i E_i^2}} = \frac{E_i}{|E|}.
\] (DH8.57)

4. [Information rudiments]
Suppose there are two fair dice. We assume that one dice is red and the other is green (that is, distinguishable). Let us record the numbers that are up in the red-green order as \((m, n)\) \((m, n \in \{1, 2, \ldots, 6\})\).

(1) To know a particular pair of numbers \((a, b)\) unambiguously what amount of information (in bits) do you need?

\((\ast)\) Actually, how many ‘yes-no’ questions do you need to pinpoint the outcome?

(2) You are told that the sum \(a + b\) is not less than 5. What is the amount of information you gain from this message?

(3) Next, you are told, one of the dice shows the face less than 3. What is the amount of information you gain? (You know the info obtained from (2) already.)

(4) Now, you are told that actually, the one of the dice in (3) is the red one. What is the amount of information of this message?

(5) Finally, you are told that face pair is actually \((2, 5)\). What is the amount of information in this final statement?

As you guess, in whatever order the information is given, the total information you gain does not depend on the actual ‘path,’ because the extent of your ignorance is a ‘state function.’

**Soln.**

(1) There are 36 distinguishable states and they are all equally probable. Therefore, the total uncertainty is \(\log_2 36 = 5.17\) bits, or the surprisal you have, when you are told, say, \((1, 1)\) actually happens, is 5.17 bits. That is, you need 5.17 bits of information to pinpoint a particular elementary event. \((\ast)\) Let us devise the ‘cleverest’ way to ask the needed questions:

(i) Divide the 36 cases into two even set of cases 18 + 18 and ask which contains the outcome. 1 bit question.

(ii) Divide the right 18 cases into two even set of cases 9 + 9 and ask which contains the outcome. 1 bit question.

(iii) Divide 9 into 4+4+1, and ask which ‘4’ contains the outcome. 1 bit question.

\((\text{iia})\) If the answer is ‘neither of them,’ we know the answer, but with probability
(iiiib) If the answer is one of ‘4’ (with probability 8/9), we must continue to ask
(iv) two questions further to pinpoint the pair in ‘4’. 2 bit questions.

Thus the expected number of the ‘questions’ is

\[ \frac{1}{9} \times 3 + \frac{8}{9} \times 5 = \frac{43}{9} < 5.17. \quad \text{(DH8.58)} \]

How come? Since the total information we need must be ‘path-independent,’ some
question(s) must not be a genuine 1 bit question.

Indeed, we exploited the discreteness of the problem and ‘cheated’ by devising
questions that are not really the 1 bit yes-no question. The question asked at 4+4+1
juncture is not an honest 1 bit question. Actually, if you wish to be as ‘fairly’ as
possible, 36 → 18 + 18 → 4 + 5 → (2, 2) or (2, 3) ··· certainly requires more than 5
questions on the average.

In this case ‘18 → 6 + 6 + 6 and 6 → 2 + 2 + 2’ is the cleverest and you need only
4 questions. In this case, the three choice question gives you \( \log_2 3 = 1.58 \) bits, so
the total information you get is \( 1 + 1.585 + 1.585 + 1 = 5.17 \) bits.

Real life lessons are that you should try to maximize the information you can get
from a single yes-no question (make cases as even as possible) and that (in real life)
an answer to a single ‘yes-no’ question could give you more than 1 bit.

(2) There is no simpler way than to list actually all the elementary states. The
following 6 states: (1,1), (1,2), (1,3), (2,1), (2,2), (3, 1) are excluded. Remaining
are 30 states, all equally probable, so \( \log_2 30 = 4.91 \) bits is the uncertainty. That is,
\[ 5.17 - 4.91 = 0.26 \] bits is the amount of information in the message.

(3) Red = 1: Green = 4, 5 or 6
Red = 2: Green = 3, 4, 5 or 6.

Therefore, there are \( 7 \times 2 = 14 \) states remaining. This uncertainty is \( \log_2 14 = 3.81 \).
We had 4.91 bits of uncertainty, so this message must have conveyed 1.1 bits.

(4) Obviously, 1 bit.

(5) There is no uncertainty remaining, so 2.81 bits (this is, needless to say, the sur-
prisal of an event of probability 1/7).

5 [Information rudiments 2]

Suppose a student D cheats in a yes-no quiz by copying the answer of an all A stu-
dent A who is correct with probability 92%. Assuming that the success rate of D is
65%, what is the information gain of D on the average by copying the solutions of
student A’s?

**Soln.**
The key point is the relation between ‘information’ and entropy’. The entropy of a state is the required information to describe the state unambiguously. Thus, decreasing entropy (importing ‘negentropy’) is adding information.

The entropy (in the Shannon sense) of student A’s answer (or his brain state)

\[ h = -(0.92 \log_2 0.92 + 0.08 \log_2 0.08) = 0.402 \text{ bits.} \]  

(DH8.59)

This means the amount of information \( 1 - h = 0.6 \) bits is needed further for student A to be perfect.

On the other the entropy of student D’s answer is

\[ h' = -(0.65 \log_2 0.65 + 0.35 \log_2 0.35) = 0.934 \text{ bits.} \]  

(DH8.60)

We need almost 1 bit (close to the entropy of randomness) of information to describe the mess of D’s brain.

Thus, his brain state is ‘improved’ (i.e., its entropy becomes closer to zero) through cheating by \( 0.934 - 0.402 = 0.532 \) bits.
Homework 8 due 9 am on Mar 26 (Tu), 2019.
Submit to compass2g

You may discuss with your friends AFTER you have made due efforts of your own
to solve the problems. I trust you. I wish you to fully understand the solutions when
you submit your homeworks (and get the full credit).

No solution without your justification will get any credit.

1 [Constant magnetic field ensemble]
“There is a lattice containing $N$ lattice sites on which non-interacting spins of $S = 1$ (in the term
$^3P$) sit. The spin Hamiltonian at each site reads
\[ H = D S_z^2. \] (DH8.61)

The system is imposed a magnetic field $B$ in the $z$-direction and the magnetic moment of of the
system is $\mu S_z$. Compute the magnetization of the system as a function of $B$ and $T$.”

This is the way a question in Qual goes, but what this tells you is that the problem
is just as:
At each site is an entity\(^{12}\) that can take $s = 0$ or $\pm 1$, whose energy is given by
\[ h = D s^2. \] (DH8.62)

The potential energy of the entity in a magnetic field $B$ (actually, the $z$-component
of $B$) is given by $U = -\mu s B$. Compute $M = \mu N \langle s \rangle$ as a function of $T$ and $B$.

(1) What is the thermodynamic coordinates of this “magnetic” system?
(2) We wish to describe the system under constant $B$ and $T$ instead of constant $M$
and $E$. What is the most convenient thermodynamic potential $\Psi$?
(3) Find the most convenient partition function $\tilde{Z}$ such that $\Psi = -k_B T \log \tilde{Z}$.
(4) Compute $\Psi$.
(5) Write down its Gibbs relation for $\Psi$. What is
\[ \left( \frac{\partial \Psi}{\partial T} \right)_B ? \] (DH8.63)

(6) Compute the magnetization $M$ as a function of the magnetic field and tempera-
ture.
(7) Can you compute the internal energy of the system?

\(^{12}\) actually the $z$-component of a spin.
2 [Information rudiments]
There are red and green tetrahedral fair dice. Thus, the outcome one simultaneous slowing gives \((m, n), (m, n \in \{1, 2, 3, 4\})\).

(1) Initially you do not know anything about the outcome. How many bits do you need to pinpoint the outcome? What is the ‘entropy’ \(h\) of your brain state?

(2) Can you devise a scheme to pinpoint the outcome on the average less than 4 ‘yes-no’ questions?

(3) You get a message that the sum \(n + m > 4\). After knowing this, what is your brain ‘entropy’? What is the information you have obtained from the message?

(4) Then, you are told that the red gives larger face value than the green. After knowing this, what is your brain ‘entropy’? What is the information you have obtained from the message?

(5) Now you are told the outcome: \((4, 3)\). What is the information you get from this final message?