Entropy in Physics: equilibrium, maximisation & minimisation principles, and the concept of free-energy

Or, why do polymers crumple and puddles evaporate?

Richard E. Spinney and Richard G. Morris

November 16, 2023

Abstract

We give an overview of the concept of entropy as applied to physical systems and illustrate how it underpins the concept of thermal equilibrium. In doing so we will be able to grasp a deeper understanding of the often-used, but maybe less well understood, notions of temperature and free energy.

INTRODUCTION AND DEFINITIONS

The concept of entropy has historically been associated with two core branches of physics: *thermodynamics* and *statistical mechanics*. Both, however, broadly sought (and indeed seek) to explain the behaviour of large aggregations of matter, originally in the context of the technology of the time such as steam engines etc. However the ultimate reach of these fields and their developments is almost completely universal. By way of orienting ourselves, we note that the title of these notes, and indeed the contents of the previous sentence have thrown the names of several concepts and fields at you so it may be worth appreciating what, in broad terms, they all mean, and how they relate to what we have seen in this series so far.

- *Thermodynamics*: Quite literally, the dynamics of heat - or more broadly energy. Energy is the absolute currency of the physical world, and so we can think of this much more broadly as attempting to find answers to the question “what will matter do?” when energy is permitted to flow around (or moved around through interventions).
- *Equilibrium*: The notion of a body at rest, i.e. a lack of motion. *Thermal* equilibrium is thus the absense of aggregate flows of heat/energy and where all thermodynamic quantities are constant in time. Often (but not always!) a system will reach thermal equilibrium when left alone for a long time.
- *Statistical Mechanics*: The study of matter, in aggregate, through probability theory, when the studied system can exist in many states.

It is through the last of these that the notion of entropy, as we have seen it so far, is introduced to the seemingly disparate question of how energy flows around and comes to rest, as it is a statistical measure, defined through a probability distribution, of such aggregations of matter. The connection here is that our understanding of a great many systems, on the level of these aggregations, *is* statistical in nature. The quantities that we observe are frequently statistical averages, of a tumultuous and chaotic microstructure we either cannot observe or cannot hope to describe. The entropy of the system is then one of these averages, playing a key role in explaining how the entire system behaves.

The form of entropy used in statistical mechanics is almost identical to that which we have seen in information theory (see *Information, data and entropy*), but there are also subtle differences, primarily concerning units.

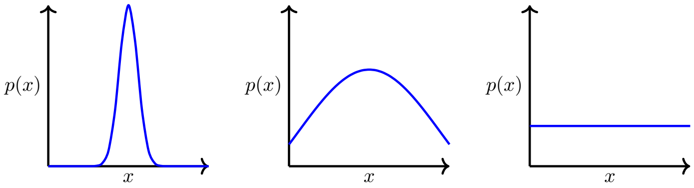
Thermodynamic entropy

The thermodynamic entropy[1]*, denoted* S*, is very closely related to the Shannon entropy,* H*. Convention dictates that* S *has units of Joules per Kelvin, and is therefore multiplied by Boltzmann's constant,* $k_B = 1.380^{-23}$ *J/K. Thus*

$$S = -k_B \int dx\, p(x) \ln p(x). \tag{1}$$

There are two basal properties of entropy that we need to get intuition for to understand how it relates to physical systems. These are

- When the probability distribution is compact, such that all the matter is closely packed or found in one place, the entropy is low. And when the matter is spread out or found across many places, the entropy is high. Indeed, the highest possible entropy probability distribution is one that is equal everywhere - where everything is as spread out as possible.



Low entropy High entropy Highest entropy

Advanced: Why does entropy increase for broader distributions?

An obvious mathematical question is *why* entropy is larger for broader distributions. It is all to do with how the combination of a probability and a negative logarithm of that probability sitting inside the integral, $-p \ln p$, vary as a simple function of a single value of p. This function (along with many other simple functions) have a very special property which is as follows

$$-(p_1 + p_2) \ln \frac{p_1 + p_2}{2} \geq -p_1 \ln p_1 - p_2 \ln p_2.$$

This property[2] is called *convexity* and is immensely important throughout thermodynamics, not just the behaviour of entropy. Try it, you will see it is valid for all probabilities (and indeed all values of $p_1, p_2 > 1$ too).

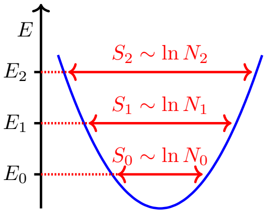
What this tells us is that by *sharing* probability between two states (i.e. by going from p_1 in state 1 and p_2 in state 2 to their average, $(p_1 + p_2)/2$, in both) the contribution $-p \ln p$ from those states will *always* go up. The entropy is just a sum of $-p \ln p$ contributions and so the entropy goes up too. This process of 'sharing' probability equally between states is just a round and about way of saying flattening the distribution.

- At thermal equilibrium, as we increase the energy in a system, the entropy also increases, and vice versa. This is broadly because as the energy of a system increases, more states become accessible to it. For example, think of a simple system such as a pendulum. At low (or indeed no) energy, it stays in one place - it has one state. But as we add energy, with increasingly large swings, we can find the pendulum in one of many more states over time.

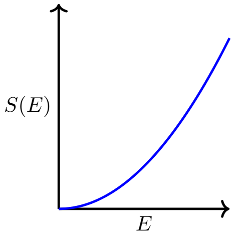
---

[1] Note, technically, this is the so-called *Gibbs entropy*.

[2] or rather a generalisation of it



Accessible states $N_2 > N_1 > N_0$



Entropy increases with energy

LAWS OF MATTER

With these concepts loosely sketched out, our next step is to briefly summarise two of the most important rules, or laws, of the aggregations of matter which we are talking about. These, of course, are the first and second laws of thermodynamics, which we will paraphrase here.

The first law

Matter and energy cannot be created or destroyed.

What this means for us: energy and matter has to go somewhere.

The second law

The entropy of a closed[3] *system cannot decrease.*

What this means for us: on long timescales it tells us where that energy/matter will go.
Remember: Between equilibrium *states, an increase in entropy accompanies an increase in energy.*

These are *axioms* of thermodynamics and we will simply be assuming them here. The discussion surrounding *why* these are, or should be, true is surprisingly deep and beyond the scope of these notes.

EQUILIBRIUM AND ENTROPY MAXIMISATION

Let us compare the second law, the fact that entropy does not decrease, with the concept of thermal equilibrium. In equilibrium all thermodynamic quantities are constant. We know that the entropy cannot decrease when in a state of equilibrium because of the second law. But what is stopping it from *increasing*? The second law certainly does not prevent that. Why has the universe reached a stationary point? The idea, crucially, is that equibrium corresponds to a *maximisation* of entropy - the entropy physically cannot get any larger. The use of 'physical' here is important - the universe can't, for example, marshall energy into existence just so it can increase entropy. We must, in this instance, obey the first law too.

This idea of maximising entropy is absolutely essential to thermodynamics and underpins what we mean by equilibrium. If the entropy can be increased it is not the equilibrium state *by definition.* Thus, the manifest tendency of systems to approach equilibrium is the realisation of an entropy *maximisation principle.*

Principle of maximum entropy

Given fixed (average) energy the entropy of a system will maximise.

Equilibrium $\Leftrightarrow$ *Maximum entropy state.*

---

[3] A closed system is one that cannot exchange energy, through heat/work/radiation or any other mechanism, or exchange matter with the rest of the universe.

Let us illustrate a simple example of entropy maximisation - we consider an ideal gas in an imaginary piston which is completely isolated from the environment such that it is closed. This idea is illustrated in Fig. (1) where we instantaneously increase the space in which atoms can move around. As time proceeds the atoms fill the space made availiable to them - or - *the probability distribution flattens out and so the entropy increases.* We know the energy is fixed because of the first law - the atoms have to go *somewhere* - and the fact that we have specified that the system is closed. The second law then tells us *where* that fixed amount of energy will go - it will spread out as much as possible as the entropy increases - remember flatter distributions have larger entropies. Note, in principle we did not need to think about *how* the gas would behave - we know the maximum entropy distribution here (it is a flat distribution *c.f. Introduction and definitions* as the energy is independent of the distribution over space for our ideal gas), and so, inevitably, that is the state the universe evolves towards.

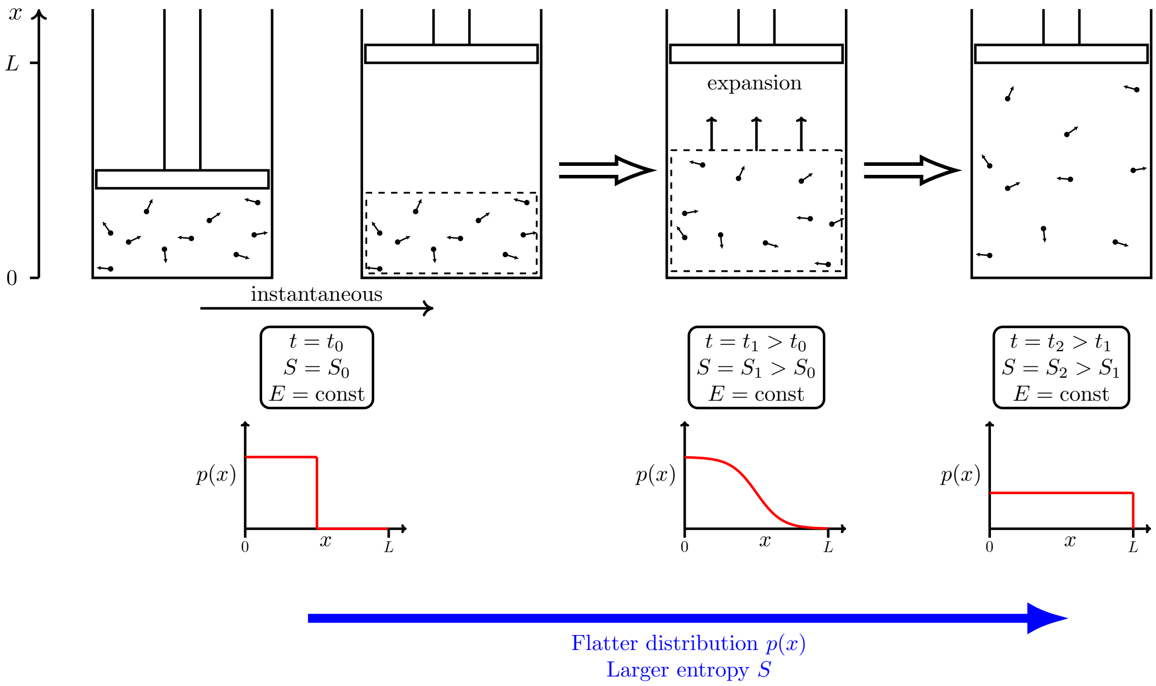


Figure 1: We consider an imaginary piston that is completely isolated from the rest of the universe containing a gas with a fixed energy $E = \text{const}$ manifesting through its velocities. The system is *closed*. If we instantaneously expand the piston, then at later times t, the gas will start to expand from a small volume to larger volumes until it occupies the full space of the piston. The probability distribution over space *broadens* and the entropy increases. All whilst the energy is a constant.

Advanced: Deriving the Boltzmann distribution using the entropy maximisation principle

The principle of entropy maximisation can be very powerful. Here we show how it is possible to derive the Boltzmann distribution using this principle alone. Of particular note here is that, unlike for the idealised closed piston, different distributions will correspond to different average energies. Dealing with this then requires some slightly advanced mathematical machinery. The idea, however, is the same - of all distributions with the set amount of energy, we choose (or rather the universe chooses) the distribution with the highest entropy.

Imagine a generic system than can exist in states x, each associated with an energy $E(x)$. If the system has a distribution over states $p(x)$ then the average energy $E(x)$ is given by $U := \langle E \rangle = \int dx\, p(x)E(x)$.

When searching for the equilibrium distribution we can then ask: what is the distribution that maximises the entropy $S = -k_B \int dx\, p(x) \ln p(x)$, when the average energy is held constant. Mathematically, this is written as:

$$p^*(x) = \underset{p(x)}{\arg\max} \left\{ -k_B \int p(x) \ln p(x)\, dx \,\middle|\, \int p(x)E(x)\, dx = U \right\}. \tag{2}$$

This is achieved by contructing the following functional[4] on $p(x)$:

$$\mathcal{F}\left[p\left(x\right)\right] = -\int p\left(x\right)\ln p\left(x\right)dx - \lambda\left(\int p(x)E(x)\,dx - U\right) \tag{3}$$

where λ is a Lagrange multiplier[5] for the constraint of constant energy. $\mathcal{F}[p(x)]$ is maximised when the following is satisfied

$$\frac{\delta\mathcal{F}[p(x)]}{\delta p\left(x\right)} = \ln p^{*}\left(x\right) + 1 + \lambda\,E\left(x\right) = 0. \tag{4}$$

The result is that

$$p^{*}\left(x\right) = C\,e^{-\frac{\lambda\,E(x)}{k_B}}, \tag{5}$$

for some constant C.

λ, as it turns out, is precisely equal to $1/T$ (see exercise **??**). In the above it responds to the fixed total energy U in the system, such that if U is larger, then λ is smaller.

Since p^{*} is a probability distribution, it must be normalised— *i.e.*, $\int p^{*}(x)\,dx = 1$. This implies that $C = 1/Z$, where

$$Z = \int e^{-\frac{E(x)}{k_B T}}\,dx, \tag{6}$$

is a very important quantity referred to as the partition function.

A classic example of the Boltzmann distribution concerns the velocities of atoms in an ideal gas. In this case we have $x = \boldsymbol{v} = \{v_x, v_y, v_z\}$ and $E(x) = E(\boldsymbol{v}) = m|\boldsymbol{v}|^2/2$ where m is the mass of a single atom. Therefore velocities are distribution according to

$$p(\boldsymbol{v}) = Z^{-1}e^{-\frac{m|\boldsymbol{v}|^2}{2k_B T}}. \tag{7}$$

From this equation one can transform the velocitities to *speeds*, $v = \sqrt{|\boldsymbol{v}|^2}$, to give the famous Maxwell-Boltzmann distribution

$$p(v) \propto v^2 e^{-\frac{mv^2}{2k_B T}}. \tag{8}$$

---

[5] A functional is something that takes a function, here our trial probability distribution $p(x)$, and gives us a number back. For instance the entropy itself is a functional.

[5] Lagrange multipliers are part of the machinery that allows us to keep the energy fixed. Specifying our constraint with an actual number would result in setting λ to some specific value too.

Dealing with the real world - the thermal environment

Now, however, is where we must deal with the practicalities of real systems. All of the above is in reference to an unspecified ‘closed system’, for which we know the second law to be valid. This is almost never the case in the real world. Whether we are studying steam engines, pistons, proteins, dna, or cells we are invariably dealing with *open* systems - systems that can exchange energy (or indeed matter) with an external environment.

And here is where themodynamics performs a book-keeping trick that has become perhaps its most recognisable feature: it separates the universe (a presumed closed system subject to the second law of thermodynamics) in two. These two components are a *system* that we care about, and an *environment* which we need to account for, but are largely unconcerned about. Moreover let us say that the system depends weakly upon the environment such that they are uncorrelated. With such assumptions we can write the total entropy of the universe in a very convenient way.

Entropy of a system plus its environment

If we have a system, with states x, and an environment, with states y, if they are uncorrelated we have $p(x, y) = p(x)p(y)$ and their mutual information is zero (See Mutual information*). Under such circumstances the entropy of the universe (the joint system x and y) decomposes as*

$$\begin{aligned} S_{\text{universe}} &= -\int dx \int dy \; p(x,y) \ln p(x,y) \\ &= -\int dx \int dy \; p(x,y) \ln p(x)p(y) \\ &= -\int dx \int dy \; p(x,y)(\ln p(x) + \ln p(y)) \\ &= -\int dx \; p(x) \ln p(x) - \int dy \; p(y) \ln p(y) \\ &= S_{\text{system}} + S_{\text{environment}} \end{aligned} \tag{9}$$

The central conceit of thermodynamics now is that our rules about entropy and conservation of energy apply to the *universe* (an ostensibly closed system), but we want to talk only about our *system*, a mere part of it.

But we can still make progress. What if we then said that our environment was very big, much much bigger than the system we care about, and that nothing interesting is happening there - it is at thermal equilibrium and always will be. We have here, somewhat crudely, described a *heat bath.* But this is very powerful. If it is always at equilibrium, its entropy has maximised - it can’t increase or decrease unless we add or take away energy. But the first law must hold. And so any energy we do add or take away must be perfectly matched by an opposite change in the system. So, let us take another look at our entropy maxmisation principle. It says that for *fixed energy* the entropy will maximise. Well, if the energy is fixed in the system, then it must be fixed in the environment (by the first law). And if it is fixed in the environment, the entropy in the environment is fixed too, because it is at equilibrium. Consequently when we maxmise the system entropy *we also maximise the entropy of the universe* and so the principle remains in tact! The question, however, remains - what does it mean to fix an energy in an open system? Or rather, how does the amount of energy in a system get determined?

To give an answer to this we need to reverse the logic of the entropy maximisation principle, in which by fixing the energy of the system we fixed the entropy of the environment. What if, instead, we fixed the entropy of the system? It turns out we have another extremisation principle, one that should be very familiar, now concerning the energy of the system

Principle of minimum energy

Given fixed entropy the energy of a system will minimise.

What has happened here? By fixing the system entropy, all changes in the entropy of the universe are due to changes in the entropy of the environment. So, the second law as applied to the universe, now is acting solely on the environment. But as the environment is at equilibrium it can only increase its entropy by increasing its energy. The first law then tells us that any increase in energy in the environment is mirrored by an equal decrease in the system. Thus, maxmising the entropy of the universe, in this situation, corresponds to minimising the energy of the system.

This is the reason behind just about every energy minimisation principle you will have seen from common day experiences where energy is lost through friction, to chemical reactions flowing down energy landscapes producing compounds with stronger bindings than their reagents. In all cases any excess energy is transferred away to the environment, where it spreads far and wide under the action of the second law, overwhelmingly likely never to return, until all excess energy has been dissipated and the system energy cannot be reduced anymore - it has been minimised.

One might ask why the principle of minimum energy is so familiar to us, why in everyday experience do we see systems minimise their energy, but not maximise their entropy. A, technically correct, but perhaps unilluminating answer is that the value of k_B is very very small. But what this means, in terms of the extremisation principles we have laid out, is that at the level of resolution that we describe macroscopic objects, the number of states they can occupy is approximately constant on the scale of the energy changes they are ever likely to experience. E.g. a ball at the top of a cliff has a similar number of states in can be in as it does at the bottom. This is *precisely* fixing the system entropy as written in the energy minimisation principle. A trivial everyday example is illustrated in Fig. 2. When a ball that experiences friction rolls down a hill it loses its kinetic energy as heat to the environment, thus minimising its potential energy. This is merely the maximisation of entropy in the environment, here being a proxy for the entropy of the universe as the entropy of the ball is fixed - we describe it with one state regardless of its potential energy[6].

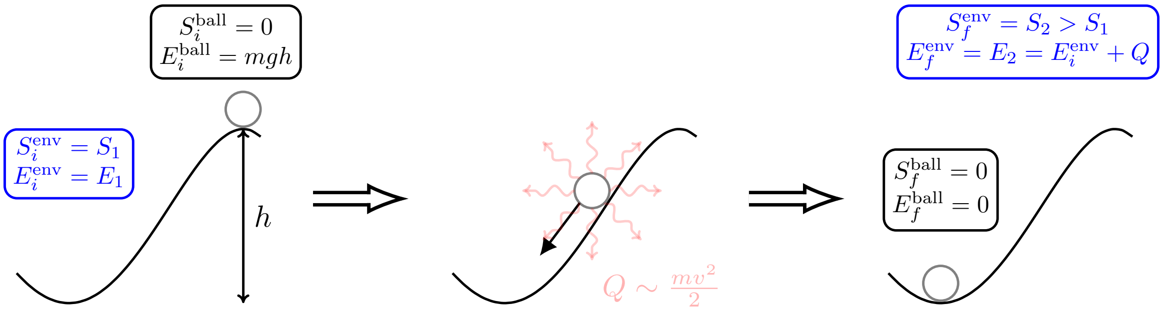


Figure 2: Entropy maximisation is behind all kinds of everyday phenomena. We can ask about why the energy of systems tends to minimise? For example, if friction is in play (as in our everyday lives), when a ball rolls down a hill, it will eventually come to rest at a local minimum or valley. This is facilitated by the fact that the ball loses its potential energy, first to kinetic energy, and then to the environment as heat. This is explainable through the second law. The entropy of the ball is fixed - it only has one state, always with probability 1. Consequently the entropy of the environment is maxmised by ‘stealing’ any excess energy from the system. In other words: all the energy being concentrated in the ball is a ‘narrower’ distribution across the universe than if the energy was distributed throughout the whole environment.

PUTTING IT ALL TOGETHER - FREE ENERGIES AND THE DEFINITION OF TEMPERATURE

So far we have seen that by assuming the first and second laws we can arrive at powerful extremisation principles which correspond to entropy maximisation of the universe being expressed solely through the system or the environment by fixing the entropy in the other. But what should a real system do when neither is fixed? Well, we go back to the second law and our fully closed system - it tries to maximise its entropy, that of the universe, by bringing the whole universe into equilibrium. But we want to express everything in terms of what we care about - the system alone. We don’t have an expression for the entropy of the universe in terms of just the system, but we know how it *varies* - it goes up with system entropy (as the system entropy is just a component of it), and down with system energy (as that requires a loss of energy from the environment). So we can say that *changes*[7] in it will

---

[6] To be crystal clear here, the entropy of a macroscopic object, such as a ball, absolutely will change as *heat* is transferred to and from it. But both this energy and entropy do not concern its *macroscopic* dynamics - the dynamics that we as observers can fully specify. In other words, the number of microscopic states a ball (for example) can be in does change as it gets hotter or colder, but it doesn’t change with height.

[7] Strictly one must be careful, some steps that follow will only be valid for *small* changes

behave as

$$\begin{aligned}\Delta S_{\text{universe}} &= \Delta S_{\text{system}} + \Delta S_{\text{environment}} \\ &= \Delta S_{\text{system}} + \alpha \Delta U_{\text{environment}} \\ &= \Delta S_{\text{system}} - \alpha \Delta U_{\text{system}},\end{aligned}$$

where we have used the first law to identify $\Delta U_{\text{environment}} = -\Delta U_{\text{system}}$. But we don't know *a priori* what α is - we don't know how much entropy increases in the environment if we increase its energy. This parameter α, as written, thus encodes the *relationship* between increases in entropy and energy in the environment. i.e. it should be understood as

$$\alpha = \frac{\Delta S_{\text{environment}}}{\Delta U_{\text{environment}}} \to \frac{dS_{\text{environment}}}{dU_{\text{environment}}}.$$

This, by *definition*, is the rigorous conception of (the inverse of) *temperature* in thermodynamics, in this case in the environment - the large heat bath where we expect to have a single well defined temperature.

Definition of temperature in thermodynamics

$$T = \frac{\partial U}{\partial S} \Leftrightarrow \frac{1}{T} = \frac{\partial S}{\partial U} \tag{10}$$

High *temperatures are characterised by* small *changes in entropy when energy is added.*
Low *temperatures are characterised by* large *changes in entropy when energy is added.*

Increases in entropy upon the increase of energy corresponds to the positivity of temperature.

We now have an expression for changes in the total entropy of the universe in terms of the system, with the entire environment characterised through a single parameter, T.

$$\Delta S_{\text{universe}} = \Delta S_{\text{system}} - \frac{\Delta U_{\text{system}}}{T}.$$

Multiplying both sides by $-T$ we find

$$-T\Delta S_{\text{universe}} = \Delta U_{\text{system}} - T\Delta S_{\text{system}}.$$

If we are to maximise the entropy of our closed system, the universe, then we must *minimise* this quantity. We recognise this quantity as a change in *free energy.*

Free energy

The change in free energy[8] *of a system in contact with an environment at temperature T is*

$$\Delta F = \Delta U - T\Delta S. \tag{11}$$

where thermodynamic quantities U, S are for the system.
This defines the bare free energy as

$$F = U - TS. \tag{12}$$

Minimising free energy is thus synonymous with maximising the entropy of the universe.

This then leads to

Principle of minimum free energy

The free energy of a system will minimise.

We can therefore appreciate why free energies are so important when describing thermal phenomena, such as processes in the cell and beyond. It is because they reflect what the system *will do* based on the balance of entropy maximisation and energy minimisation in the system which is, ultimately, a consequence of the second law applied to both, i.e. the universe, simultaneously.

Mesostates

So far we our usage of the word *state* has been referring to microscopic configurations of our systems, often referred to a *microstates*. Another notion of state in thermodynamics concerns that of the macroscopic behaviour - a *macrostate* which we can think of as synonymous with our distribution $p(x)$ over the microstates, thus giving us values for the mean energy U and entropy S. There are, however, times when neither is suitable for what we are trying to achieve. A common scenario where this might arise is where we want to talk about sets of microstates as single comparable units. For instance we might be concerned with whether a protein is folded or unfolded, or whether a dna strand is bound or not, each of which having some non zero probability of occuring, without having to concern ourselves with every single configuration of every bond in the macromolecule in question etc. Such aggregations of microstates are often called *mesostates* and dealing with them requires a little bit of care, but is worthwhile, as it is the form of free energy that is widely reported within biological contexts. The idea of how microstates are grouped together into mesostates is illustrated in Fig. (3).

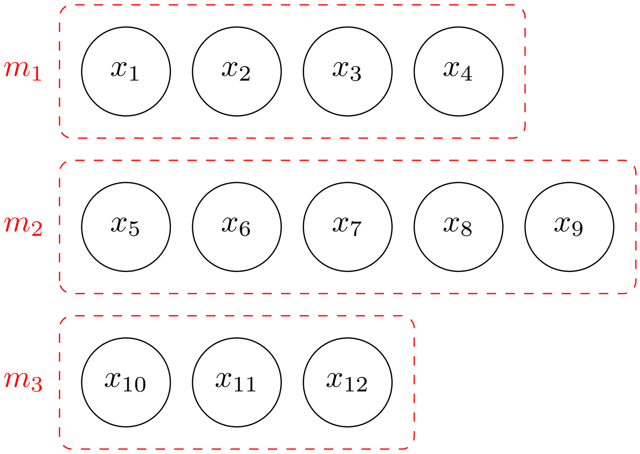


Figure 3: A system of 12 microstates x, is aggregated into 3 mesostates m. Each microstate x has energy $E(x)$, whilst a mean energy can be associated with a mesostate. For instance $U_{m_1} = \sum_{i=1}^{4} p(x_i)E(x_i)$. The probability of being in a mesostate equals the sum of probability in its microstates. E.g. $p(m_2) = \sum_{i=5}^{9} p(x_i)$.

Putting the treatment of all mesostates together is slightly involved, and described below for completeness, but the two single most important things we need to know about their use is as follows

- We treat each mesostate as a small system in its own right, each attempting to maximise the joint entropy of the environment and the mesostate. Indeed they all have their own average energy, entropy, and free energy which may differ between mesostates. We achieve this by replacing every occurence of the probability of the microstate x, $p(x)$, with the conditional probability of that microstate given that we are only considering mesostate m, $p(x|m)$. We therefore have mestostate quantities:

---

[8] Note, this is the *Helmoltz* free energy

Mesostate energies, entropies, and free energies

For a given mesostate m we have entropy, mean energy, and free energy

$$\begin{aligned} S_m &= -k_B \int_{x\in m} dx\; p(x|m)\ln p(x|m) \\ U_m &= \int_{x\in m} dx\; p(x|m)E(x) \\ F_m &= U_m - TS_m. \end{aligned}$$

Note: $p(x|m) = p(x,m)/p(m) = p(x)/p(m)$, *and notation* $\int_{x\in m}$ *means 'integrate over all microstates* x *in mesostate* m*'.*

- The probability of a mesostate being occupied follows an adapted Boltzmann formula where energy is replaced by free energy of the mesostate. I.e.

Probability of a mesostate

A mesostate m*, with free energy* F_m*, is occupied in proportion to:*

$$p(m) \propto e^{-F_m/k_BT}$$

Crucially, whenever, in some micro-biological context, a broad configuration or state is associated with a free energy, it is this kind of meso-state free energy that is being reported. Examples are myriad, but might include free energies of a folded or unfolded configuration of a protein, or specific binding configuration between strands of dna.

We must emphasise, the high level understanding is identical to what we have seen before. In attempting to maximise the entropy of the universe, states that have lower energy and higher entropy are preferred. Thus the favourable mesostate out of two with the same entropy, but different mean energies is that with the lower energy. And the the favourable mesostate out of two which have the same energy, but different entropies, is that with the higher entropy. In crudest terms, mesostates are exponentially favoured according to their mean energy, but are also exponentially favoured according to how many ways there are to exist in that mesostate.

Advanced: Probability of a mesostate and relationship between global free energy and meso-state free energies

When dealing with mesostates in more detail, we first must set out some broad rules and assumptions. We insist that every individual microstate is found in one, and only one, mesostate m. This means that all original integrals of the form $\int dx$ are now entirely equivalent to the sum and integral $\sum_m \int_{x\in m}$.

We can demonstrate that the probability of a mesostate goes as e^{-F_m/k_BT} as follows

$$\begin{aligned} F_m &= \int_{x\in m} dx\; p(x|m)\left(E(x) + k_BT\ln p(x|m)\right) \\ &= \int_{x\in m} dx\; p(x|m)\left(E(x) + k_BT\ln\frac{p(x)}{p(m)}\right) \\ &= \int_{x\in m} dx\; p(x|m)\left(E(x) + k_BT\left(\ln p(x) - \ln p(m)\right)\right) \\ &= \int_{x\in m} dx\; p(x|m)\left(E(x) + k_BT\left(\ln Z^{-1}e^{-E(x)/k_BT} - \ln p(m)\right)\right) \\ &= \int_{x\in m} dx\; p(x|m)\left(k_BT\ln Z^{-1} - k_BT\ln p(m)\right) \\ &= k_BT\ln Z^{-1} - k_BT\ln p(m) \\ \Rightarrow p(m) &= Z^{-1}e^{-F_m/k_BT}. \end{aligned} \tag{13}$$

Note, we have used $p(x) = p(x,m) = p(x|m)/p(m)$ and therefore $p(x|m) = p(x)/p(m)$ on the second line.

We can then see how F_m relates to the total system free energy by rearranging the sum/integral used in its constituent components, viz.

$$\begin{aligned}
F &= U - TS \\
&= \int dx\ p(x)E(x) + k_B T \int dx\ p(x) \ln p(x) \\
&= \sum_m \int_{x\in m} dx\ p(x)E(x) + k_B T \int_{x\in m} dx\ p(x) \ln p(x) \\
&= \sum_m p(m) \int_{x\in m} dx\ p(x|m)E(x) + k_B T \sum_m p(m) \int_{x\in m} dx\ p(x|m) \ln p(x) \\
&= \sum_m p(m)U_m + k_B T \sum_m p(m) \int_{x\in m} dx\ p(x|m) \ln p(x) \\
&= \sum_m p(m)U_m + k_B T \sum_m p(m) \int_{x\in m} dx\ p(x|m) \ln p(x|m)p(m) \\
&= \sum_m p(m)U_m + k_B T \sum_m p(m) \int_{x\in m} dx\ p(x|m) \ln p(x|m) + k_B T \sum_m p(m) \ln p(m) \\
&= \sum_m p(m)U_m - k_B T \sum_m p(m)S_m - TS_I \\
&= \sum_m p(m)(U_m - k_B T S_m) - TS_I \\
&= \sum_m p(m)F_m - TS_I \\
&= \langle F_m \rangle - TS_I
\end{aligned} \tag{14}$$

where

$$S_I = -k_B \sum_m p(m) \ln p(m) \tag{15}$$

is a residual contribution to the total entropy *over the mesostates*, and thus free energy, which cannot be associated with any individual mesostate.