§10. Thermodynamic Equilibria and Stability

In §9, I proved that the stat. mech. T and S are the same as the thermodynamical T and S.

This was nice to know but, strictly speaking, unnecessary. Instead, I could have *defined*

$$S = k_B S_G \quad \text{and} \quad T = \frac{1}{k_B \beta}$$

(with a historical factor k_B to show respect for Ludwig Boltzmann)

and then *constructed all of thermodynamics as a consequence of Stat. Mech.*, without ever having to go through all those heat engines, Carnot cycles etc.

[so the way it has been presented to you is *chronological*: thermodynamics was worked out in the 19th century before Stat. Mech. in early 20th; *logically* we don't really *need a separate construction of thermo.*, except as an example of how to set up an empirical theory of physical phenomena whose microscopic nature one does not understand]

We saw on p.82 how the entire *thermodynamic calculus* (based on eq.(10) + specific expression for F or Z) comes out immediately out of S.M. This will be applied to monatomic ideal gas in §11 and other things in PART IV.

But we also need to deal with some *conceptual* issues:

- conditions for th. *equilibrium* (in our new language this means state of max S subject to measurable constraints)
  (constant $\underset{\text{thermal eq.}}{\underset{\uparrow}{T}}$, constant $\underset{\text{mechanical eq.}}{\underset{\uparrow}{p}}$, no $\underset{\text{dynamical eq.}}{\underset{\uparrow}{\text{velocity}}}$ gradients)
- its *stability* (if $\delta S = 0$, how do we know it's a max, not min?)
- *2nd law* — tricky issue, to be discussed later on.

10.1 Additivity of entropy

Consider 2 systems:

① microstates α, energy levels $E^{(1)}_\alpha$, probabilities $p^{(1)}_\alpha$

② microstates α', energy levels $E^{(2)}_{\alpha'}$, probabilities $p^{(2)}_{\alpha'}$

Put them together into a composite system, but in such a way that the 2 sub-systems are in "loose" thermal contact, viz., the microstates of the 2 sub-systems are independent (in QM language, eigenstates of the composite system are products of eigenstates of its parts) [works, e.g., for gases or fluids, but not for solids, where states are collective] [see Binney's QM §6.1]

Then, for the composite system, the microstates are enumerated by

$$(\alpha,\alpha') \Leftrightarrow \text{energy levels } E_{\alpha\alpha'} = E^{(1)}_\alpha + E^{(2)}_{\alpha'}$$

$$\text{probabilities } p_{\alpha\alpha'} = p^{(1)}_\alpha \cdot p^{(2)}_{\alpha'} \quad (\alpha,\alpha' \text{ independent!})$$

Then the Gibbs entropy

$$S_G = -\sum_{\alpha\alpha'} p_{\alpha\alpha'} \ln p_{\alpha\alpha'} = -\sum_{\alpha\alpha'} p^{(1)}_\alpha p^{(2)}_{\alpha'} \ln\left(p^{(1)}_\alpha p^{(2)}_{\alpha'}\right) =$$

$$= -\sum_\alpha p^{(1)}_\alpha \ln p^{(1)}_\alpha \underbrace{\sum_{\alpha'} p^{(2)}_{\alpha'}}_{=1} - \sum_{\alpha'} p^{(2)}_{\alpha'} \ln p^{(2)}_{\alpha'} \underbrace{\sum_\alpha p^{(1)}_\alpha}_{=1}$$

$$= S^{(1)}_G + S^{(2)}_G \qquad (1)$$

entropy is additive

(in equilibrium, $S = k_B S_G$, so $S = S^{(1)} + S^{(2)}$)

So is energy:

$$U = \sum_{\alpha\alpha'} E_{\alpha\alpha'} p_{\alpha\alpha'} = \sum_{\alpha\alpha'} E^{(1)}_\alpha p^{(1)}_\alpha p^{(2)}_{\alpha'} + \sum_{\alpha\alpha'} E^{(2)}_{\alpha'} p^{(1)}_\alpha p^{(2)}_{\alpha'}$$

$$= U^{(1)} + U^{(2)} \qquad (2)$$

Note that $E_{\alpha\alpha'} = E^{(1)}_{\alpha} + E^{(2)}_{\alpha'}$ implies that <u>partition functions multiply</u>

$$Z(\beta) = \sum_{\alpha\alpha'} e^{-\beta E_{\alpha\alpha'}} = \sum_{\alpha\alpha'} e^{-\beta\left(E^{(1)}_{\alpha} + E^{(2)}_{\alpha'}\right)} =$$

$$= \left(\sum_{\alpha} e^{-\beta E^{(1)}_{\alpha}}\right)\left(\sum_{\alpha'} e^{-\beta E^{(2)}_{\alpha'}}\right) = Z^{(1)}_{(\beta)} Z^{(2)}_{(\beta)} \qquad (3)$$

and, therefore, <u>in equilibrium</u>,

$$p_{\alpha\alpha'} = \frac{e^{-\beta E_{\alpha\alpha'}}}{Z} = \frac{e^{-\beta E^{(1)}_{\alpha}}}{Z^{(1)}} \frac{e^{-\beta E^{(2)}_{\alpha'}}}{Z^{(2)}} = p^{(1)}_{\alpha} p^{(2)}_{\alpha'} \qquad (4)$$

and also

$$F = -k_B T \ln Z = F^{(1)} + F^{(2)} \qquad (5)$$

$$S = -\left(\frac{\partial F}{\partial T}\right)_V = S^{(1)} + S^{(2)} \qquad (6)$$

$$U = -\frac{\partial \ln Z}{\partial \beta} = U^{(1)} + U^{(2)} \qquad (7)$$

10.2 Thermal Equilibrium

Put 2 systems (in loose thermal contact) together, otherwise isolated.

Then

$$U = U_1 + U_2 = \text{const}$$

$$S = S_1 + S_2 \rightarrow \max$$

(has to be maximized, as we know, to find what the new eq. state is)

So,

$$dS = dS_1 + dS_2 =$$

$$= \frac{\partial S_1}{\partial U_1} dU_1 + \frac{\partial S_2}{\partial U_2} dU_2$$

but $dU = dU_1 + dU_2 = 0$ so $dU_2 = -dU_1$ *)

$$= \left(\frac{\partial S_1}{\partial U_1} - \frac{\partial S_2}{\partial U_2}\right) dU_1 = 0 \quad \text{(maximize)} \qquad (8)$$

---

*) Note: This is a Lagrange multiplier calculation: we are maximising $S_1 + S_2$ subject to $U_1 + U_2 = U$, so $S_1 + S_2 - \lambda(U_1 + U_2 - U) \rightarrow \max$ gives $\left(\frac{\partial S_1}{\partial U_1} - \lambda\right)dU_1 + \left(\frac{\partial S_2}{\partial U_2} - \lambda\right)dU_2 + (U_1 + U_2 - U)\,d\lambda = 0 \Rightarrow \frac{\partial S_1}{\partial U_1} = \frac{\partial S_2}{\partial U_2} = \lambda = 1/T$ etc.

Recall that (eq.(10) p.82)

NB: This formula is only valid for eq. states. Using it here means we are assuming both systems 1&2 are in eq. and so is the composite system 12

$$dS = \frac{1}{T}dU + \frac{P}{T}dV \qquad (9)$$

$$\Downarrow$$

$$\boxed{\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_V} \qquad (10)$$

← and const any other "external" parameter, if any

So, eq. (8) is

$$dS = \left(\frac{1}{T_1} - \frac{1}{T_2}\right)dU_1 = 0 \Rightarrow \boxed{T_1 = T_2}$$

Thus, the temperatures must be equal (thermal equilibrium)
(the condition for everything to have been in equilibrium, as assumed at the outset)

NB: The validity of (10) did not depend on any identification of T and S with the same quantities in empirical thermodynamics, it is true for the stat. mech. temperature (Lagrange multiplier) and stat. mech. entropy (measure of uncertainty in $\{p_\alpha\}$). So, the above shows that the S.M. temperature is a sensible definition of temperature: it is something that is equal, in equilibrium, between two systems in loose energy contact with each other.

This property allows us to make a thermometer and hence introduce a temperature scale.

[Reminder: In thermodynamics T was introduced either via 0th law (2 system are in th. eq. with a 3rd ⇒ then they are in th. eq. with each other) or as the universal integrating factor of δQ (Kelvin)]

↑ see discussion on pp. 81 & 83-84 (showing those def's are equivalent)

10.3 Heat bath and the canonical ensemble

At this point, it is worth pausing and discussing what the canonical distribution actually describes physically. Recall that it was set up by stipulating that probabilities of the microstates would be maximally unbiased subject only to conspiring to give some fixed (measurable) value of avg. energy U. The result was Gibbs distribution

$$p_\alpha = \frac{e^{-\beta E_\alpha}}{Z(\beta)},$$

where β was to be determined as a f-n of U via

$$U = -\frac{\partial \ln Z}{\partial \beta} \Rightarrow \beta = \frac{1}{k_B T} = \beta(U, \dots)$$

But both the structure of the theory that has emerged (*implicit* eqn for β) and the experience/anticipation of the kinds of questions we are likely to be interested in suggest that in fact it is much preferable to have $U = U(T, \dots)$, with T as "input parameter."

Why is this preferable? Because T of a system is often known by virtue of the system being in contact with surroundings (= *heat reservoir*, *heat bath*) whose T is fixed — usually because the system is small compared to the h. bath and so can draw from or give up to the heat bath arbitrary amounts of energy w/o affecting the T of the h. bath very much.

So this indeed is what the canonical ensemble describes: microstates of a system in thermal contact with a heat bath of fixed T.

In equilibrium, $T_{system} = T_{bath}$, as we proved on p. 89.

<u>Note</u>: One can explicitly construct Gibbs distribution on this basis if one starts from a (fictional) "closed system" with equal probabilities for all microstates ("<u>microcanonical ensemble</u>") and then considers a small part of it. I will discuss this later on.

10.4 Mechanical and dynamical equilibria

So far, we have only treated energy as a variable quantity (exchangeable between systems / or, equiv., between the system and the heat bath it is in), while treating the volume as a fixed external parameter and also assuming implicitly that everything was static (no velocity).

Let us generalise this consideration and consider some number of systems, indexed by i, and each having total energy $\mathcal{E}_i$, mass m_i, velocity $\vec{u}_i$, centre of mass at position $\vec{r}_i$ and volume V_i.
Join them all together (in a "loose contact", as explained in 10.1, so their microstates remain independent) and allow them to exchange energy, momentum, angular momentum and also push on each other ("exchange volume").
If we now isolate them and confine them in a box of fixed volume, the equilibrium state of the combined system must be the state of max entropy subject to conservation laws:

$\sum_i \mathcal{E}_i = \mathcal{E}_0$ total energy (NB: $\mathcal{E}_i$ includes both internal and kinetic energy)

$$\sum_i m_i \vec{u}_i = \vec{P} \quad \text{total momentum}$$

$$\sum_i m_i \vec{r}_i \times \vec{u}_i = \vec{L} \quad \text{total angular momentum} \qquad (11)$$

$$\sum_i V_i = V \quad \text{total volume of the box}$$

The entropy of each subsystem can depend only on the internal energy, not that of the macroscopic motion: this is because of Galilean invariance (probabilities of microstates $\{p_\alpha\}$ are the same in any inertial frame). Therefore,

$$S_i = S_i\Big(\underbrace{\mathcal{E}_i - \frac{m_i u_i^2}{2}}_{\substack{\shortparallel \\ U_i \text{ internal energy}}}, V_i\Big) \qquad (12)$$

Maximise:

$$\sum_i S_i - \lambda\Big(\sum_i \mathcal{E}_i - \mathcal{E}\Big) - \vec{a}\cdot\Big(\sum_i m_i \vec{u}_i - \vec{P}\Big) - \vec{b}\cdot\Big(\sum_i m_i \vec{r}_i \times \vec{u}_i - \vec{L}\Big)$$
$$-\sigma\Big(\sum_i V_i - V\Big) \to \max$$

Lagrange multipliers

Lagrange multipliers (vectors because each of these cons. laws is a set of 3 scalar eqns)

Maximise this wrt $\mathcal{E}_i, \vec{v}_i, V_i$:

$$\frac{\partial S_i}{\partial \mathcal{E}_i} - \lambda = 0 \qquad (13) \text{ thermal eq.}$$

$$\frac{\partial S}{\partial \vec{u}_i} - m_i \vec{a} - m_i \vec{b}\times\vec{r}_i = 0 \qquad (14) \text{ dynamical eq.}$$

$$\frac{\partial S}{\partial V_i} - \sigma = 0 \qquad (15) \text{ mechanical eq.}$$

Eq. (13): $\dfrac{\partial S_i}{\partial \mathcal{E}_i} = \dfrac{\partial S_i}{\partial U_i} = \dfrac{1}{T_i}$ [see eq. (10)] $\Rightarrow \boxed{\dfrac{1}{T_i} = \lambda \equiv \dfrac{1}{T}}$ (16)

All temperatures equal to the same L. multiplier and so equal to each other. This is simply the generalisation to many systems of the result of 10.2 (thermal eq.)

Eq. (15): $\frac{\partial S_i}{\partial V_i} = \frac{p_i}{T_i}$ ← pressure! because $dS = \frac{1}{T}dU + \frac{P}{T}dV$

But all $T_i = T$, so $$\boxed{\frac{p_i}{T} = \sigma \equiv \frac{P}{T}} \quad (17)$$ from (15)

Note that for i.g.s, this L. multiplier is ~ density

All pressures are the same — mechanical equilibrium

(everything is in pressure balance, otherwise volumes would expand or shrink)

Eq. (14): $\frac{\partial S_i}{\partial \vec{u}_i} = \frac{\partial S_i}{\partial \mathcal{E}_i}\left(-m_i\vec{u}_i\right) = -\frac{m_i\vec{u}_i}{T_i} = -\frac{m_i\vec{u}_i}{T}$

↖ all $T_i = T$

Then from (14),

$$\boxed{\vec{u}_i = -\underbrace{\vec{a}T}_{\text{overall velocity}} - \underbrace{T\vec{b}}_{\text{overall angular velocity}}\times\vec{r}_i \equiv \vec{u} + \vec{\Omega}\times\vec{r}_i} \quad (18)$$

Thus, the only macroscopic motion in equilibrium can be an overall constant motion in some direction plus rigid-body rotation of the whole system — dynamical equilibrium.

The meaning of these results is simply that in equilibrium, there can't be any temperature or pressure gradients or any internal macroscopic motions (velocity gradients).

SM does not tell us how this is achieved, but we know, e.g., from our experience in KT that ∇T's and $\nabla\vec{u}$'s will relax via thermal diffusivity and viscosity, respectively.

Note: 1) In practice, mech. equilibrium is often achieved faster than the thermal and dynamical one, at least in incompressible systems: pressure imbalances ⇒ uncompensated forces ⇒ macroscopic motions, etc.

2) These arguments are generalised in an obvious way to non-PV systems

3) Another type of equilibrium we might have considered is particle equilibrium – by allowing our subsystems to exchange particles, subject to the overall conservation of the total #. This leads to the equality of chemical potentials of all systems – another Lagrange multiplier, which we will introduce in PART V when we study "open systems".

4) In considering quantities other than energy as measurable constraints, we went beyond the canonical ensemble – indeed, further ensembles can be constructed, where besides mean energy, other things are fixed: e.g. mean particle # ("grand canonical ensemble"), mean volume ("pressure ensemble"), mean angular momentum ("rotational ensemble"), etc.

5) There is no separate ensemble for the momentum. Indeed if we consider non-rotating systems, eq. (18) says $\vec{u}_i = \vec{u}$. We can always go to the frame where $\vec{u} = 0$ and so nothing moves.

L3 ended here.

10.5 Stability

How do we know that when we extremized $\mathcal{S}$, the solution we found was a max, not a min? This is equivalent to asking whether the equilibria we find are stable.

This amounts to calculating second derivatives of the entropy:

- thermal equilibrium: eq. (13), (16)

$$\frac{\partial^2 S_i}{\partial \mathcal{E}_i^2} = \frac{\partial \lambda}{\partial \mathcal{E}_i} \underset{\text{eq.(16)}}{\underset{\uparrow}{=}} \frac{\partial}{\partial \mathcal{E}_i}\frac{1}{T} = -\frac{1}{T^2}\frac{\partial T}{\partial \mathcal{E}_i} < 0$$

cond'n of stability

$$\Updownarrow$$

$$\boxed{\frac{\partial \mathcal{E}_i}{\partial T} = C_V > 0} \qquad (19)$$

That heat capacity is always positive can be proven in about three lines directly by calculating $C_V = \left(\frac{\partial U}{\partial T}\right)_V$ from $U = -\frac{\partial \ln Z}{\partial \beta}$ and sub-ing the explicit Gibbs expression for Z (Ex.).

It turns out
$$C_V = \frac{\langle (E_\alpha - U)^2 \rangle}{k_B T^2} > 0 . \qquad (20)$$

For simplicity, let us only consider the case of a system with fixed volume and no rotation ($\vec{\Omega} = 0$).

- dynamical equilibrium :

$$\frac{\partial^2 S_i}{\partial u_i^\mu \partial u_i^\nu} = \frac{\partial}{\partial u_i^\mu}\left[\frac{\partial S_i}{\partial \mathcal{E}_i}\left(-m_i u_i^\nu\right)\right] = -m_i \underbrace{\frac{\partial S_i}{\partial \mathcal{E}_i}}_{=\frac{1}{T}} \delta_{\mu\nu} - m_i u_i^\nu \frac{\partial}{\partial u_i^\mu}\frac{\partial S_i}{\partial \mathcal{E}_i}$$

vector indices

$$= -\frac{m_i}{T}\delta_{\mu\nu} + m_i^2 \underbrace{u_i^\mu u_i^\nu}_{=0} \underbrace{\frac{\partial^2 S_i}{\partial \mathcal{E}_i^2}}_{=-\frac{1}{C_V T^2}} = -\frac{m_i}{T}\delta_{\mu\nu} < 0$$

because can always go to frame with $\vec{u}=0$

cond'n of stability

$$\Downarrow$$

$$\boxed{T>0} \qquad (20)$$

Thus, we have proven that temperature must be positive !

(systems with negative temperature are unstable)

Another, more qualitative way of arguing this is as follows:

$$S = \sum_i S_i\left(\mathcal{E}_i - \frac{m_i u_i^2}{2}\right)$$

If $\frac{\partial S_i}{\partial \mathcal{E}_i} = \frac{1}{T} < 0$, then all S_i's would be maximized by decreasing their argument as much as possible, i.e., increasing all u_i's subject to $\sum_i u_i = 0$. This means parts of the system would fly in opposite directions (everything would blow up).

[Although the prohibition on negative T can in some sense be relaxed if bits of the system are not allowed to move,] and if the system has a max. possible energy.

Finally, note a similar argument for *the positivity of pressure*: $p = T\left(\frac{\partial S}{\partial V}\right)_U$.

If this is <0, entropy in a (closed) system can increase if V goes down, i.e., the system will shrink to nothing. In contrast, when $p>0$, entropy increases as V increases (system expands) – but this is checked by walls or whatever maintains fixed V.

This argument* militates strongly against negative pressure, but it is not completely prohibitive – negative pressures can exist (although usually in *metastable states*, to be discussed in PART VI on real gases) – this happens for example when cavities form or substances separate from walls etc.