Virial Theorem (Classical systems)

Consider a stable system of discrete particles, bound by potential forces (forces characterized exclusively by potential difference). The virial theorem provides a general relation between the total kinetic energy and the total potential energy of the system. The theorem states

$$\langle K.E\rangle = -\frac{1}{2}\sum_{i=1}^{N}\langle \vec{F}_i \cdot \vec{r}_i\rangle$$

Here K.E stands for the total kinetic energy of N particles, $\vec{F}_i$ represents the force on the ith particle, which is located at position $\vec{r}_i$, and the angular brackets represent the time average. The significance of the virial theorem is that it allows the average total kinetic energy to be calculated even for very complicated systems that defy an exact solution, such as those considered in statistical mechanics.

Consider

$$\langle x_i \frac{\partial H}{\partial x_j}\rangle = \frac{\int dq_i\, dp_i\ x_i \frac{\partial H}{\partial x_j}\ e^{-\beta H}}{\int dq_i\, dp_i\ \ e^{-\beta H}}$$

where x_i and x_j are any of the 6N generalized coordinates, for example, p or q.

$$\int dq_i\, dp_i\ x_i \frac{\partial H}{\partial x_j}\ e^{-\beta H} = -\frac{1}{\beta}\int dq_i\, dp_i\ x_i \frac{\partial e^{-\beta H}}{\partial x_j}$$

Integrating by parts we get

$$\int dq_i\, dp_i\ x_i \frac{\partial e^{-\beta H}}{\partial x_j} = \int'_{\substack{integrating \\ over\ all\ coordinates \\ except\ x_j}} dq_i\, dp_i\ x_i e^{-\beta H}\Bigg]_{x_j^{(1)}}^{x_j^{(2)}} - \int dq_i\, dp_i\ \frac{\partial x_i}{\partial x_j}\ e^{-\beta H}$$

Here $x_j^{(1)}$ and $x_j^{(2)}$ are the extremal values of x_j. The boundary integral vanishes because H becomes infinite at the extremal values of any coordinates. For example, if x_j is momentum p, then the extremal values of $p = \pm\infty$ and $H = \frac{p^2}{2\,m} \to \infty$. For example, if x_j is spatial coordinate q, then the extremal values are at boundary of system, where the potential energy confining the particle to the volume V becomes infinite. This implies,

$$\int dq_i\, dp_i\ x_i \frac{\partial H}{\partial x_j}\ e^{-\beta H} = \frac{1}{\beta}\int dq_i\, dp_i\ \frac{\partial x_i}{\partial x_j}\ e^{-\beta H}$$

But $\frac{\partial x_i}{\partial x_j} = \delta_{ij}$. Thus, we get

$$\langle x_i \frac{\partial H}{\partial x_j}\rangle = \frac{1}{\beta}\ \ \delta_{ij}\ \frac{\int dq_i\, dp_i\ \ e^{-\beta H}}{\int dq_i\, dp_i\ \ e^{-\beta H}}$$

$$\langle x_i \frac{\partial H}{\partial x_j} \rangle = k_B T \ \ \delta_{ij}$$

Is the called the Virial theorem.

If $x_i = x_i = p_i$ then

$$\langle p_i \frac{\partial H}{\partial p_i} \rangle = \langle p_i \dot{q}_i \rangle = k_B T$$

If $x_i = x_i = q_i$ then

$$\langle q_i \frac{\partial H}{\partial q_i} \rangle = -\langle q_i \dot{p}_i \rangle = k_B T$$

Where we have used Hamiltonian equations of motion $\frac{\partial H}{\partial p_i} = \dot{q}_i$ and $\frac{\partial H}{\partial q_i} = -\dot{p}_i$.

If we consider 3N degrees of freedom, then

$$\langle \sum_{i=1}^{3N} p_i \dot{q}_i \rangle = 3N\, k_B T$$

$$-\langle \sum_{i=1}^{3N} q_i \dot{p}_i \rangle = 3N\, k_B T$$

In simple terms,

$$2\langle Total\ K.E \rangle = -\langle Total\ P.E \rangle$$

Here, $3N\, k_B T = 2\, \frac{3N\, k_B T}{2} = 2\, Total\ K.E$

Paramagnetism (Classical systems)

Consider N classical spins (paramagnetic system), and ignore the interaction between the spins. When an external magnetic field is applied, these spins tend to align in the direction of the field. For this case, we can write down the Hamiltonian

$$H = -\sum_{i=1}^{N} \vec{\mu}_i \cdot \vec{h} = -\mu\, h \sum_{i=1}^{N} \cos \alpha_i$$

where $\vec{\mu}_i$ is magnetic moment of spin i, $|\vec{\mu}_i| = \mu$, α_i is the angle of $\vec{\mu}_i$ with respect to $\vec{h}$.
As the spins are non-interacting, the partition function for the system (N-spins) is

$$Q_N = (Q_1)^N$$

Here we need not include $N!$ term because the spins are distinguishable. We can imagine that each spin sits at a fixed position in space and so can be distinguishable from any other spin.

For the consider system,

$$Q_1 = \sum_{\substack{\alpha \\ (All\ possible\ angles)}} e^{\beta\mu h \cos\alpha}$$

Here, α may be oriented in an arbitrary direction with respect to the applied magnetic field. Thus, one needs to consider all possible configurations of α.

$$Q_1 = \int_0^{2\pi} d\phi \int_0^{\pi} d\theta\ \sin\theta\ e^{\beta\mu h \cos\theta}$$

It is As there is no ϕ dependence, the above integral simplifies into,

$$Q_1 = 2\pi \int_0^{\pi} d\theta\ \sin\theta\ e^{\beta\mu h \cos\theta}$$

Take $\cos\theta = x$. Implies, $\sin\theta\, d\theta = -dx$. The corresponding limits are $x = 1$ (for $\theta = 0$) and $x = -1$ (for $\theta = \pi$). Thus,

$$\int_0^{\pi} d\theta\ \sin\theta\ e^{\beta\mu h \cos\theta} = \int_{-1}^{1} dx\ e^{\beta\mu h\, x}$$

$$= \frac{1}{\beta\mu h}\left[e^{\beta\mu h\, x}\right]_{-1}^{1} = \frac{2}{\beta\mu h}\left[\frac{e^{\beta\mu h} - e^{-\beta\mu h}}{2}\right]$$

$$= \frac{2}{\beta\mu h}\frac{\sinh(\beta\mu h)}{\beta\mu h}$$

So, we get

$$Q_1 = 2\pi \int_0^{\pi} d\theta\ \sin\theta\ e^{\beta\mu h \cos\theta} = 4\pi\,\frac{\sinh(\beta\mu h)}{\beta\mu h}$$

$$Q_1 = 4\pi\,\frac{\sinh(\beta\mu h)}{\beta\mu h}$$

Now, we can calculate the average total magnetization $\vec{M}$ oriented along $\vec{h}$ direction. If we choose $\vec{h} = h\,\hat{z}$ (along z direction), then

$$M_z = N\,\langle \mu \cos\alpha \rangle$$

Here $\langle \mu \cos\theta \rangle$ is the projection of one spin along the direction of the magnetic field.

$$M_z = N\,\langle \mu\cos\theta \rangle = N\,\frac{\sum_\alpha \mu\cos\theta\ e^{\beta\mu h\cos\alpha}}{\sum_\alpha e^{\beta\mu h\cos\alpha}}$$

$$M_z = N\,\frac{\frac{1}{\beta}\frac{\partial}{\partial h}\left(\sum_\alpha e^{\beta\mu h\cos\alpha}\right)}{\sum_\alpha e^{\beta\mu h\cos\alpha}}$$

$$M_z = \frac{N}{\beta}\,\frac{\frac{\partial}{\partial h}(Q_1)}{Q_1}$$

Here, $Q_1 = 4\pi \frac{\sinh(\beta\mu h)}{\beta\mu h}$. Thus, we get

$$M_z = \frac{N}{\beta} \frac{4\pi \left[\frac{\cosh(\beta\mu h)}{h} - \frac{\sinh(\beta\mu h)}{\beta\mu h^2}\right]}{4\pi \frac{\sinh(\beta\mu h)}{\beta\mu h}}$$

$$M_z = N\mu \left[\coth(\beta\mu h) - \frac{1}{\beta\mu h}\right]$$

The average magnetization is

$$\langle\mu_z\rangle = \frac{M_z}{N} = \mu \left[\coth(\beta\mu h) - \frac{1}{\beta\mu h}\right]$$

The term in the bracket is called the Langevin function. If we take $y = \beta\mu h$, then in the $y \to \infty$ limit, $\coth(\beta\mu h) \to 1$. So, the term in the bracket tends to one. Implies,

$$\langle\mu_z\rangle = \mu$$

It means the magnetization per spin is constant which is μ.

However, the other limit, i.e., $y \to 0$, is more interesting. In this limit, the Langevin function can be simplified as

$$\coth(y) - \frac{1}{y} = \frac{\cosh(y)}{\sinh(y)} - \frac{1}{y}$$

$$= \frac{1 + \frac{y^2}{2!}}{y + \frac{y^3}{3!}} - \frac{1}{y} = \frac{1 + \frac{y^2}{2}}{y\left(1 + \frac{y^2}{6}\right)} - \frac{1}{y}$$

$$= \frac{\left(1 + \frac{y^2}{2}\right)\left(1 - \frac{y^2}{6}\right)}{y} - \frac{1}{y} = \frac{1 + \frac{y^2}{2} - \frac{y^2}{6}}{y} - \frac{1}{y} = \frac{y}{3}$$

This implies,

$$\langle\mu_z\rangle = \frac{\mu^2 \beta h}{3} = \frac{\mu^2 h}{3k_B T}$$

The net magnetization is

$$M_z = \frac{N \mu^2 h}{3k_B T}$$

The magnetic susceptibility

$$\chi = \lim_{h\to 0} \frac{\partial M_z}{\partial h} = \frac{N \mu^2}{3k_B T}$$

$$\chi \propto \frac{1}{T}$$

This is called Curie's law of paramagnetism. The magnetic susceptibility of a paramagnetic material grows linearly with $1/T$.

Entropy and information

The entropy S can be expressed in terms of the probability P as

$$S = -k_B \sum_i P_i \ln P_i$$

where P_i is the probability to be in state i. This relation holds for both canonical and microcanonical ensemble.

Claude Shannon (1948) turned this relation backwards, in developing a close relation between entropy and information theory. Consider a system with states labelled by i, and P_i is the probability for the system to be in state i. We want to define a measure of how disordered the distribution P_i is. Call the disorder measure S (it will turn out to be the entropy).

$$S = -k \sum_i P_i \ln P_i$$

where k is the proportionality constant.

The above relation states that the bigger (smaller) S is, the more (less) disordered the system is, the less (more) information we have about the probable state of the system. Here, S satisfies the following conditions.

1. If $p_j = \begin{cases} 1 & j = i \\ 0 & j \neq i \end{cases}$ Then the state of the system is exactly known to be i.

This should have $S = 0$ as there is no uncertainty, no disorder.

2. For equally likely p_i, i.e. all p_i = 1/N for N states, the system is maximally disordered. It means S is maximum possible value for all possible N state distributions.

3. S should be additive over independently random systems.