

Statistical Physics Review Notes

Matthew Ho

Carnegie Mellon University 33-765, Spring 2018

Taught by: Markus Deserno

Reference: Swendsen 2012 [Sw]

1 Entropy

1.1 The Classical Ideal Gas

[Sw 11-15]

An 'ideal' gas differs from a 'real' gas in that its particles do not interact. Most real gases at low densities act as near ideal gases, because the average distance between particles is relatively high.

Our model of a classical gas consists of N particles with well-defined positions and momenta in a specified volume. We assume that each component of the position and momenta vectors can be treated as independent degrees of freedom. We use this to create an abstract $6N$ -dimensional phase space.

$$q = \{\mathbf{r}_i | i = 1, \dots, N\} = \{q_j | j = 1, \dots, 3N\} ; \quad p = \{\mathbf{p}_i | i = 1, \dots, N\} = \{p_j | j = 1, \dots, 3N\}$$
$$\{q, p\} = \{q_j, p_j | j = 1, \dots, 3N\} ; \quad E = \sum_{i=1}^N \frac{|\mathbf{p}_i|^2}{2m} \quad (1)$$

For an ideal gas, there are no interactions between particles, so the potential energy is zero.

Particles will be regarded as distinguishable when the exchange of two particles result in a different microscopic state. Classically, this is equivalent to saying every point in phase space represents a different microscopic state.

In 1877, Boltzmann defined entropy to be proportional to the logarithm of the 'probability' of a state.

$$S = k_B \log W \quad (2)$$

where S is the entropy of the macrostate, k_B is Boltzmann's constant ($1.38 * 10^{-23} J/K$), and W is the number of possible microstates which would exhibit the macrostate. We assume that the positions and momenta of particles are independent quantities such that their contributions to the total entropy are additive.

1.2 Discrete Probability Theory

[Sw 16-39]

1.2.1 Probability

Frequentists define probability as the asymptotic frequency of success in the limit of an infinite number of trials.

$$p = \lim_{N \rightarrow \infty} \frac{N_s}{N} \quad (3)$$

where N is the number of trials in an experiment, and N_s is the number of successes. The other camp of belief, the Bayesians, believe that this definition is unsatisfactory, as we will never be able to run infinite trials of an experiment and thus we will never be able to measure a probability.

Bayesians describe probability as a function of a person's knowledge at the time of the trial. The probability that a Bayesian might report might change as he gathers more evidence. This belief provides an appropriate way to describe what we learn from experiments.

Sw describes 'model probability' to differentiate from absolute probability, in the frequentist or Bayesian sense. A model probability is an assumption as to what the frequency of success would be for an infinite number of trials. Statistical mechanics is based on simple assumptions, expressed as model probabilities, that lead to predictions in agreement with experiment.

We can define a set of elementary events A with associated probabilities $P(a_j)$

$$A = \{a_j | j = 1, \dots, N_A\} ; \quad 0 \leq P(a_j) \leq 1 ; \quad \sum_{j=1}^{N_A} P(a_j) = \sum_a P(a) = 1 \quad (4)$$

1.2.2 Multiple Random Variables

We can extend this to describe a situation with two or more sets of random events. Including event A and a similar event B will give rise to a joint probability $P(a, b)$

$$B = \{b_k | k = 1, \dots, N_B\} ; \quad 0 \leq P(a, b) \leq 1 ; \quad \sum_a \sum_b P(a, b) = 1 \quad (5)$$

We define marginal probability and conditional probability, respectively, as follows

$$P_A(a) = \sum_b P(a, b) ; \quad P(a|b) = \frac{P(a, b)}{P_B(b)} \quad (6)$$

From these follows the famous Bayes' theorem. Bayes' theorem is significant because it gives us a mechanism for adjusting our predictions based on prior evidence.

$$P(a|b) = \frac{P(b|a) P_A(a)}{P_B(b)} \quad (7)$$

Two variables can be classified as independent if their joint probability distribution is equal to the product of their marginal probabilities.

$$P(a, b) = P_A(a) P_B(b) ; \quad P(a|b) = P_A(a) \quad (8)$$

Three or more independent variables can have mutual or pairwise independence.

1. Pairwise independence means that the marginal distribution of any pair of random variables can be written as the product of the marginal distributions of the individual random variables.
2. Mutual independence means that the marginal distribution of any subset of random variables can be written as the product of the constituent marginal distributions.

Mutual independence implies pairwise independence, but the converse is not always true.

1.2.3 Functions of Random Variables

Given an arbitrary random variable A , define a numerical function on the set of elementary events $F = \{F(a_j)|1, \dots, N_A\}$. The probability distribution of F is given by

$$P_F(f) = \sum_a \delta_{f,F(a)} P_A(a) ; \quad \delta_{x,y} = \begin{cases} 1, & x = y \\ 0, & x \neq y \end{cases} \quad (9)$$

where $\delta_{x,y}$ is the Kronecker delta. For an arbitrary function $G(x,y)$ on two random variables X and Y , one can describe its probability distribution as

$$P_G(g) = \sum_x \sum_y \delta_{g,G(x,y)} P(x,y) \quad (10)$$

1.2.4 Derived quantities

The mean is defined as the first moment of the function. The mean and the n -th moment of a function $F(A)$ defined on the random variable A are given by

$$\langle F \rangle \equiv \sum_a F(a) P_A(a) ; \quad \langle F^n \rangle \equiv \sum_a F(a)^n P_A(a) \quad (11)$$

The n -th moment is a linear operation. The n -th central moment is defined by

$$\langle (F - \langle F \rangle)^n \rangle \equiv \sum_a (F(a) - \langle F \rangle)^n P_A(a) \quad (12)$$

The variance can be written

$$\sigma_F^2 = \langle (F - \langle F \rangle)^2 \rangle ; \quad \sigma_F^2 \equiv \langle F^2 \rangle - \langle F \rangle^2 \quad (13)$$

One method of describing how two random variables is through their correlation and covariance.

$$\text{Cov}(X, Y) = \langle XY \rangle - \langle X \rangle \langle Y \rangle ; \quad c_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} \quad (14)$$

If F and G are independent random variables, the correlation $c_{FG} = 0$. However, the converse is not always true.

Given a set of random numbers $\{F_j|j = 1, \dots, N\}$, we are often interested in their sum.

$$S = \sum_{j=1}^N F_j ; \quad \langle S \rangle = \sum_{j=1}^N \langle F_j \rangle \quad (15)$$

If the random numbers are pairwise independent,

$$\sigma_S^2 = \sum_{j=1}^N \sigma_j^2 \quad (16)$$

Furthermore, if the random numbers have the same mean and variance,

$$\langle S \rangle = N \langle F \rangle ; \quad \sigma_S^2 = N \sigma^2 \quad (17)$$

Sets of random, independent variables with the same probability distributions are denoted as 'independent and identically distributed' (iid). These have the property of a relative standard deviation which decreases with the square root of the number of variables.

$$\frac{\sigma_S}{\langle S \rangle} = \frac{\sigma}{\langle F \rangle \sqrt{N}} \quad (18)$$

1.2.5 Binomial Distribution

The binomial distribution is representative of a case of N iid random numbers, $\{F_j\}$ that can take on the value of 1 or 0 with probabilities p or $1 - p$, respectively. The mean and variance of F are

$$\langle F \rangle = p ; \quad \sigma_F^2 = p(1 - p) \quad (19)$$

The mean and variance of the sum of N iid variables, S , follows as $\langle S \rangle = N \langle F \rangle$ and $\sigma_S^2 = N \sigma_F^2$. The explicit probabilities are given by the binomial distribution

$$P(s) = P(n|N) = \binom{N}{n} p^n (1 - p)^{(N-n)} ; \quad \binom{N}{n} = \frac{N!}{n!(N - n)!} \quad (20)$$

where n is the number of variables taking the value 1.

1.2.6 Gaussian functions

A general, normalized Gaussian function is given by

$$g(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x - x_0)^2}{2\sigma^2}\right\}, \quad \langle x \rangle = x_0 = x_{max} ; \quad \langle (x - x_0)^2 \rangle = \sigma^2 \quad (21)$$

For sufficiently large N and a fixed value of p , the Gaussian function acts as a good approximation to the binomial distribution.

$$x_0 = pN ; \quad \sigma^2 = p(1 - p)N \quad (22)$$

Here's a useful procedure for Gaussian integrals

$$\begin{aligned} G &= \int_{-\infty}^{\infty} \exp\{-ax^2\} dx = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \exp\{-y^2\} dy \\ G^2 &= \frac{1}{a} \int_{-\infty}^{\infty} \exp\{-(x^2 + y^2)\} dx dy = \frac{1}{a} \int_0^{\infty} \exp\{-r^2\} 2\pi r dr = \frac{\pi}{a} \\ G &= \sqrt{\frac{\pi}{a}} \end{aligned} \quad (23)$$

1.2.7 Stirling's Approximation

For large values of N , $N!$ can be difficult to calculate, even computationally. One can produce a rough approximation for $N!$ for high N . The simplest form of Stirling's approximation is

$$N! \approx N^N \exp\{1 - N\} ; \quad \log N! \approx N \log N - N \quad (24)$$

A more stringent version is

$$N! \approx e^{-N} N^N \sqrt{2\pi N} \quad (25)$$

An even better version of Stirling's approximation is given by Gosper

$$N! \approx e^{-N} N^N \sqrt{(2N + \frac{1}{3})\pi} \quad (26)$$

However, for most purposes, the simplest approximation is sufficient.

1.2.8 Characteristic Functions

The Fourier transform $\tilde{p}(k)$ of a probability density $p(x)$ is called the characteristic function.

$$\tilde{p}(k) = \langle e^{ikx} \rangle = \int dx p(x) e^{ikx} ; \quad p(x) = \frac{1}{2\pi} \int dk \tilde{p}(k) e^{-ikx} \quad (27)$$

Some unique characteristics of the Fourier transform are

$$\mu_n = i^{-n} \left[\frac{\partial^n}{\partial k^n} \tilde{p}_X(k) \right]_{k=0} ; \quad \tilde{p}_{aX}(k) = \tilde{p}_X(ak) ; \quad \tilde{p}_{X+Y}(k) = \tilde{p}_X(k) \tilde{p}_Y(k) \quad (28)$$

1.3 The Classical Ideal Gas: Configurational Entropy

[Sw 40-46]

Our assumption that the positions and momenta of particles in an ideal gas are independent allows us to consider each separately.

$$P(q, p) = P_q(q) P_p(p) ; \quad S(E, V, N) = S_q(V, N) + S_p(E, N) \quad (29)$$

1.3.1 Distribution of Particles between Two Subsystems

Consider a composite system of two boxes containing a total of N distinguishable, non-interacting particles. The two boxes have volumes V_A and V_B . The total number of particles N and box volume $V = V_A + V_B$ are held constant. We can assume that the positions are both independent of the momenta and of each other. We'll also assume that a given particle is equally likely to be anywhere in the composite system, such that the probability of it being in box A is V_A/V .

$$P_q(q) = \prod_{j=1}^N P_1(\mathbf{r}_j) ; \quad P(q \in A) = \frac{V_A}{V} \quad (30)$$

We see that the probability distribution of the particle distribution between the two boxes can be written with the binomial distribution.

$$P(N_A|N) = \binom{N}{N_A} \left(\frac{V_A}{V} \right)^{N_A} \left(1 - \left(\frac{V_A}{V} \right) \right)^{N-N_A} ; \quad P(N_A, N_B) = \frac{N!}{N_A! N_B!} \left(\frac{V_A}{V} \right)^{N_A} \left(\frac{V_B}{V} \right)^{N_B} \quad (31)$$

Some consequences of the binomial distribution are that the average values for N_A and N_B are simply related to the volume fraction of their containers.

$$\langle N_A \rangle = N \left(\frac{V_A}{V} \right) ; \quad \langle N_B \rangle = N \left(\frac{V_B}{V} \right) ; \quad \sigma_{N_A}^2 = \langle N_A \rangle \left(\frac{V_B}{V} \right) \quad (32)$$

Sw makes a big deal about the distinction between N_A and $\langle N_A \rangle$. N_A is a property of the system which fluctuates with time. $\langle N_A \rangle$ is a description of the system and is time-independent.

1.3.2 Probability and Entropy

We will introduce a new function $\Omega_q(N, V)$ to help describe the probability distribution of the two-box system.

$$\Omega_q(N, V) = \frac{V^N}{N!} ; \quad P(N_A, N_B) = \frac{\Omega_q(N_A, V_A)\Omega_q(N_B, V_B)}{\Omega_q(N, V)} \quad (33)$$

Noting that log is a monotonic function of its argument, it will become convenient to describe $P(N_A, N_B)$ in terms of its logarithm.

$$\ln P(N_A, N_B) = \ln \Omega_q(N_A, V_A) + \ln \Omega_q(N_B, V_B) - \ln \Omega_q(N, V) \quad (34)$$

We will define another function called the configurational entropy

$$S_q(N, V) = k \ln \Omega_q(N, V) + kXN \approx kN \left[\ln \left(\frac{V}{N} \right) + X \right] \quad (35)$$

where k and X are arbitrary constants.

Noting that $N_A + N_B = N$,

$$S_{q,tot}(N_A, V_A, N_B, V_B) = k \ln P(N_A, N_B) + S_q(N, V) = S_q(N_A, V_A) + S_q(N_B, V_B) \quad (36)$$

Sw then shows that the maximum of this sum yields the location of the equilibrium values $\langle N_A \rangle$ and $\langle N_B \rangle$ that we derived in the previous section.

1.4 Continuous Random Numbers

[Sw 47-61]

We will define an extension of the probability theory discussed previously to address continuous random numbers. Instead of taking on discrete values, continuous random variables can take on any value within a specified interval.

1.4.1 Probability densities

The distribution of a random variable's output is given by its probability density function (pdf), $P(x)$. The probability of a continuous random variable X producing a value in the interval $[a, b]$ is given by

$$P([a, b]) = \int_a^b P(x) dx ; \quad P(\Omega) = \int_{\Omega} P(x) dx = 1 \quad (37)$$

where Ω is the full range over which the pdf is defined. The marginal and conditional probabilities are analogous to the definition for discrete numbers.

$$P_x(x) = \int_{-\infty}^{\infty} P(x, y) dy ; \quad P(y|x) = \frac{P(x, y)}{P_x(x)} \quad (38)$$

Independence still suggests that $P(x, y) = P_x(x)P_y(y)$. The averages, moments, and central moments are also analogous to the discrete case.

$$\langle x \rangle = \int_{-\infty}^{\infty} xP(x) dx ; \quad \langle x^n \rangle = \int_{-\infty}^{\infty} x^n P(x) dx ; \quad \langle (x - \langle x \rangle)^n \rangle = \int_{-\infty}^{\infty} (x - \langle x \rangle)^n P(x) dx \quad (39)$$

1.4.2 Jensen's inequality

Jensen's inequality is stated for a convex function $f(x)$ and a continuous random variable X .

$$\langle f(X) \rangle \geq f(\langle X \rangle) \quad (40)$$

1.4.3 Dirac Delta Functions

The Dirac delta function $\delta(x)$ is analogous to the Kronecker delta function from the discrete case.

$$\delta(x) = \lim_{\epsilon \rightarrow 0} \begin{cases} 0, & x < -\epsilon \\ \frac{1}{2\epsilon}, & -\epsilon \leq x \leq \epsilon \\ 0, & x > \epsilon \end{cases} \quad (41)$$

Integrals over the Dirac delta have unique properties

$$\int_a^b \delta(x) dx = 1 ; \quad \int_a^b \delta(cx) dx = \frac{1}{|c|}; \quad \int_{-\infty}^{\infty} f(x) \delta(c(x - x_0)) dx = \frac{f(x_0)}{|c|} \quad (42)$$

where $a < 0 < b$. We can generalize the last equation to allow the delta function to have an arbitrary argument $g(x)$. If $g(x)$ has zeros at points $\{x_j | j = 1, \dots, n\}$, and $g'(x)$ is the derivative of $g(x)$,

$$\int_{-\infty}^{\infty} f(x) \delta(g(x)) dx = \sum_{j=1}^n \frac{f(x_j)}{|g'(x_j)|} \quad (43)$$

1.4.4 Transformations of Continuous Random Variables

Given two continuous random variables, x and y , along with their joint pdf, $P(x, y)$, the pdf of a new random variable, s , defined by $s = f(x, y)$ is

$$P(s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(x, y) \delta(s - f(x, y)) dx dy \quad (44)$$

1.5 The Classical Ideal Gas: Energy-Dependence of Entropy

[Sw 62-70]

1.5.1 Distribution of energy between two subsystems

Consider a composite system of two subsystems A and B filled with classical, non-interacting particles N_A and N_B . The two subsystems can exchange energy (through a diathermal wall), but not particles. The total energy of each subsystem is given by

$$E_\alpha = \sum_{j=1}^{N_\alpha} \frac{|\mathbf{p}_{\alpha,j}|^2}{2m} \quad (45)$$

where α is A or B . The total energy $E = E_A + E_B$ and number of particles $N = N_A + N_B$ remain constant. We make the assumption that the probability distribution in momentum space is constant and uniform, conditioned on the total energy. Using the rules for transforming continuous variables, we write $P(E_A, E_B)$ as the

$$P(E_A, E_B) = P(E_A|E) = \frac{\int_{-\infty}^{\infty} \delta\left(E_A - \sum_{j=1}^{N_A} \frac{|\mathbf{p}_{A,j}|^2}{2m}\right) dp_A * \int_{-\infty}^{\infty} \delta\left(E_B - \sum_{j=1}^{N_B} \frac{|\mathbf{p}_{B,j}|^2}{2m}\right) dp_B}{\int_{-\infty}^{\infty} \delta\left(E - \sum_{j=1}^N \frac{|\mathbf{p}_j|^2}{2m}\right) dp} \quad (46)$$

Defining a function Ω_E , we can write this in a similar form to the positional probability.

$$\Omega_E(E_\alpha, N_\alpha) = \int_{-\infty}^{\infty} \delta\left(E_\alpha - \sum_{j=1}^{N_\alpha} \frac{|\mathbf{p}_{\alpha,j}|^2}{2m}\right) dp_\alpha; \quad P(E_A, E_B) = \frac{\Omega_E(E_A, N_A) \Omega_E(E_B, N_B)}{\Omega_E(E, N)} \quad (47)$$

This is a $3N$ dimensional integral, one for each axis of each of N particles. To evaluate Ω_E , we notice that the delta function makes the integrand vanish everywhere except when

$$2mE = \sum_{j=1}^N |\mathbf{p}_j|^2 \quad (48)$$

We then transform Ω_E from a $3N$ dimensional integral to a one dimensional integral

$$\Omega_E(E, N) = \int_0^\infty S_n p^{3N-1} \delta\left(E - \frac{p^2}{2m}\right) dp; \quad p^2 = \sum_{j=1}^N |\mathbf{p}_j|^2 \quad (49)$$

where S_n is the surface area of an n dimensional sphere with unit radius. This integration comes out to

$$\Omega_E(E, N) = S_n m (2mE)^{\frac{3N-1}{2}} \quad (50)$$

Sw derives the volume and surface area of an n -dimensional sphere

$$V_n = C_n r^n; \quad S_n = n C_n r^{n-1}; \quad C_n = \frac{\pi^{n/2}}{(n/2)!} \quad (51)$$

Thus, the exact expression for Ω_E is

$$\Omega_E(E, N) = \frac{3N \pi^{3N/2}}{(3N/2)!} m (2mE)^{(3N-1)/2} \quad (52)$$

$$\ln \Omega_E(E, N) = \frac{3N \pi^{3N/2}}{(3N/2)!} m (2mE)^{(3N-1)/2} \approx N \left[\frac{3}{2} \ln \left(\frac{E}{N} \right) + X \right] \quad (53)$$

Combining this with the pdf from before, the energy-dependence of the pdf is

$$P(E_A, E_B) \propto (E_A)^{(3N_A-1)/2} (E_B)^{(3N_B-1)/2} \quad (54)$$

Including $E = E_A + E_B$ and the fact that N is large, this has a maximum at

$$E_{A,max} = \left(\frac{3N_A - 1}{3N - 2} \right) E \approx \left(\frac{E}{N} \right) N_A \quad (55)$$

suggesting that the total energy is evenly distributed over all particles. The variance is given by

$$\sigma_{E_A}^2 = \frac{\langle E_A \rangle^2}{N} \left(\frac{2N_B}{3N_A} \right) \quad (56)$$

1.6 Classical Gases: Ideal and Otherwise

[Sw 71-87]

Now, we assumed at the start that the contributions to the entropy from the position and momenta of particles were independent. Therefore, we can simply sum the contributions of entropy due to these two degrees of freedom.

$$S(E, V, N) = k_B N \left[\ln \left(\frac{3N\pi^{3N/2}}{(3N/2)!} \right) + \ln \left(\frac{V^N}{N!} \right) + X' \right] \quad (57)$$

Under Stirling's approximation, we return the entropy of the classical ideal gas.

$$S(E, V, N) = k_B N \left[\frac{3}{2} \ln \left(\frac{E}{N} \right) + \ln \left(\frac{V}{N} \right) + X \right] ; \quad X = \frac{3}{2} \ln \left(\frac{4\pi m}{3h^2} \right) + \frac{5}{2} \quad (58)$$

The constant X is arbitrary under classical mechanics, but this is the traditional value chosen.

1.6.1 Equilibrium Conditions

We can now show how maximizing entropy leads to useful equilibrium conditions for two subsystems.

$$\begin{aligned} \frac{\partial S_A}{\partial E_A} &= \frac{\partial S_A}{\partial E_A} ; & \frac{\partial S}{\partial E} &= \frac{3k_B}{2E} & \rightarrow & \frac{E_A}{N_A} = \frac{E_B}{N_B} \\ \frac{\partial S_A}{\partial N_A} &= \frac{\partial S_A}{\partial N_A} ; & \frac{\partial S}{\partial N} &= k_B \left[\frac{3}{2} \ln \left(\frac{E}{N} \right) + \ln \left(\frac{V}{N} \right) + X \right] & \rightarrow & \frac{N_A}{V_A} = \frac{N_B}{V_B} \\ \frac{\partial S_A}{\partial V_A} &= \frac{\partial S_A}{\partial V_A} ; & \frac{\partial S}{\partial V} &= \frac{k_B N}{V} & \rightarrow & \frac{N_A}{V_A} = \frac{N_B}{V_B} \end{aligned} \quad (59)$$

1.6.2 Indistinguishable Particles

Sw derives the pdf for two subsystems A and B filled with N identical, non-interacting particles. The two subsystems have fixed volumes V_A and V_B and can freely exchange particles. The pdf derivation [Sw 76-78] isn't as important as the conclusion.

$$P(N_A, N_B) = \frac{N!}{V^N} \left(\frac{V_A^{N_A}}{N_A!} \right) \left(\frac{V_B^{N_B}}{N_B!} \right) \quad (60)$$

Sw concludes that this expression is identical to the distinguishable case, suggesting that the entropy of a classical gas is the same for distinguishable and indistinguishable particles.

1.6.3 Interacting particles

Real gases have interacting particles. Typical interatomic potentials are short-ranged, but can have significant consequences. We will consider the same situation as before, of two boxes A and B . We write the Hamiltonian of the system as a combination of self and inter-subsystem interaction energies.

$$\begin{aligned} H(q, p) &= H_A(q_A, p_A) + H_B(q_B, p_B) + H_{AB}(q_A, q_B) \\ H_\alpha(q_\alpha, p_\alpha) &= \sum_{j=1}^{3N_\alpha} \frac{|\mathbf{p}_{j,\alpha}|^2}{2m} + \sum_{j=1}^{3N_\alpha} \sum_{i=1, i>j}^{3N_\alpha} \phi(\mathbf{r}_{i,\alpha}, \mathbf{r}_{j,\alpha}) ; & H_{AB}(q_A, q_B) &= \sum_{j=1}^{3N_A} \sum_{i=1}^{3N_B} \phi(\mathbf{r}_{i,A}, \mathbf{r}_{j,B}) \end{aligned} \quad (61)$$

We assume that the inter-subsystem interaction term, H_{AB} can be ignored because of the small distance range of interactions. This allows us to construct a total entropy of the form

$$S_{tot}(E_A, V_A, N_A; E_B, V_B, N_B) = S_A(E_A, V_A, N_A) + S_B(E_B, V_B, N_B)$$

$$S_\alpha(E_\alpha, V_\alpha, N_\alpha) = k_B \ln \left[\frac{1}{h^{3N_\alpha} N_\alpha!} \int dq_\alpha \int dp_\alpha \delta(E_\alpha - H_\alpha) \right] \quad (62)$$

1.6.4 Second and Zeroth Law of Thermodynamics

The Second Law of Thermodynamics states that a system will tend towards maximum entropy at unconstrained equilibrium. All other non-equilibrium states must have lower total entropy. Maximizing total entropy typically leads to the following equilibrium relations

$$\frac{\partial S_A}{\partial E_A} = \frac{\partial S_B}{\partial E_B}; \quad \frac{\partial S_A}{\partial V_A} = \frac{\partial S_B}{\partial V_B}; \quad \frac{\partial S_A}{\partial N_A} = \frac{\partial S_B}{\partial N_B} \quad (63)$$

The Zeroth Law of Thermodynamics states that if two systems are each in equilibrium with a third system, they must also be in equilibrium with each other.

$$\frac{\partial S_A}{\partial E_A} = \frac{\partial S_B}{\partial E_B} = \frac{\partial S_C}{\partial E_C} \quad (64)$$

1.7 Temperature, Pressure, Chemical Potential, and All That

[Sw 88-98]

Temperature, described in the classical sense, can be summarized by the ideal gas law.

$$PV = Nk_B T \quad (65)$$

1.7.1 Derivation of the Ideal Gas Law

Sw derives the Maxwell-Boltzmann probability density for the momentum of a single particle (Sw 90-92). He does this by finding the marginal probability density of particle 1 from the full pdf $P(p, q)$ shown above. Note the equation below is correct; the book has an error.

$$P(\mathbf{p}_1) = \left(\frac{\beta}{2\pi m} \right)^{3/2} \exp \left\{ -\beta \frac{|\mathbf{p}_1|^2}{2m} \right\}; \quad \beta \equiv \frac{\partial}{\partial E} \ln \Omega(E, V, N) \quad (66)$$

Because the three components of momentum are independent, the probability density of a single component can be found by integrating out the other two.

$$P(p_{1,x}) = \sqrt{\frac{\beta}{2\pi m}} \exp \left\{ -\beta \frac{p_{1,x}^2}{2m} \right\} \quad (67)$$

Consider a flat portion of the wall of a container of an ideal gas with area A . Assume that the wall's normal is parallel to the x-axis, such that only the p_x component of momentum contributes to the pressure. The average force on the wall during a time-period Δt is given by.

$$F \Delta t = \int \Delta p_x P(p_x, \Delta t) dp_x \quad (68)$$

We'll assume that particles will only hit the wall if they are within a distance $\Delta t p_x/m$ of the wall. Therefore, the volume of particles hit by the wall is $A\Delta t p_x/m$. Assuming elastic collisions, $\Delta p_x = 2p_x$. Including the Maxwell-Boltzmann pdf in 1-D and integrating over all momenta,

$$P = \frac{F}{A} = \frac{2N}{Vm} \sqrt{\frac{\beta}{2\pi m}} \int_0^\infty \exp\left\{-\beta \frac{p_x^2}{2m}\right\} p_x^2 dp_x = \frac{N}{V\beta} \quad (69)$$

This leads to the usual gas law if we set

$$\beta = \frac{\partial}{\partial E} \ln \Omega(E, V, N) = \frac{1}{k_B T} \quad (70)$$

1.7.2 The Entropy and ...

From the explicit expression for the entropy of the classical ideal gas $S(E, V, N)$, we can find the derivative of the entropy with respect to several quantities.

$$\left(\frac{\partial S}{\partial V}\right)_{E,N} = \frac{P}{T}; \quad \left(\frac{\partial S}{\partial E}\right)_{V,N} = \frac{1}{T}; \quad \left(\frac{\partial S}{\partial N}\right)_{E,V} = -\frac{\mu}{T} \quad (71)$$

The three equations of state for an ideal gas are given by

$$PV = Nk_B T; \quad E = \frac{3}{2} Nk_B T$$

$$\mu = -k_B T \left[\frac{3}{2} \ln \left(\frac{3}{2} k_B T \right) + \ln \left(\frac{V}{N} \right) + X - \frac{5}{2} \right] \quad (72)$$

where μ is the chemical potential of the system.

1.7.3 The Differential Form of the Fundamental Relation

Using the derivatives above, we can write the differential form of the fundamental relation as

$$dS = \left(\frac{1}{T}\right) dE + \left(\frac{P}{T}\right) dV - \left(\frac{\mu}{T}\right) dN \quad (73)$$

2 Thermodynamics

2.1 The Postulates and Laws of Thermodynamics

[Sw 101-108]

2.1.1 Microscopic and Macroscopic States

A microscopic state is a property of a system of particles. In classical mechanics, a microscopic state is characterized by specific values of the positions and momenta of every particle, a point in phase space. The

microscopic state is not experimentally accessible to macroscopic measurements. A macroscopic state is a description of a thermodynamic system based on macroscopic measurements. A macroscopic state may be consistent with an infinite amount of microscopic states.

2.1.2 Postulates and Laws of Thermodynamics

The postulates are as follows

1. There exist equilibrium states of a macroscopic system that are characterized uniquely by a small number of extensive variables.
 - Extensive variables are quantities that provide a measure of the size of the system. They include energy U , volume V , and number of particles N . By contrast, intensive variables are independent of the size of the system, and include temperature T , pressure P , and chemical potential μ .
2. The values assumed by the extensive parameters of an isolated composite system in the absence of an internal constraint are those that maximize the entropy over the set of all constrained macroscopic states. ($\Delta S \geq 0$)
3. The entropy of a composite system is additive over the constituent subsystems.
 - This assumes that interactions between constituent subsystems are negligible.
4. The entropy is a monotonically increasing function of the energy for equilibrium values of the energy.
 - This postulate implies that temperature is positive.
5. The entropy is a continuous and differentiable function of the extensive parameters.
 - This postulate breaks down when there exists some discontinuous instability or phase transition in the system.

6. The entropy is an extensive function of the extensive variables.

$$S(\lambda U, \lambda V, \lambda N) = \lambda S(U, V, N) \quad (74)$$

- 'Extensivity' is the property that the entropy is directly proportional to the size of the system. This means that the entropy is assumed to be a homogeneous, first-order function of the extensive variables. Extensivity is not valid for all systems!

The laws are as follows

0. If two systems are each in equilibrium with a third system, they are also in equilibrium with each other.
1. Heat is a form of energy, and energy is conserved.
2. After the release of a constraint in a closed system, the entropy of the system never decreases.
3. The entropy of any quantum mechanical system goes to a constant as the temperature goes to zero.

Lastly, there is a universally recognized convention for the absolute value of entropy. This convention implies that

$$\lim_{T \rightarrow 0} S(T) \geq 0 \quad (75)$$

2.2 Perturbations of Thermodynamic State Functions

[Sw 109-115]

We can apply the First Law to small changes in thermodynamic quantities

$$dU = dQ + dW \quad (76)$$

2.2.1 Inexact Differentials

Inexact differentials dF can be integrated along any path, but will not vanish on a closed loop. For example $dF = x dx + x dy$. If a differential is exact, we can write it in the form

$$dF = \left(\frac{\partial F}{\partial x} \right) dx + \left(\frac{\partial F}{\partial y} \right) dy \quad (77)$$

for some function $F(x, y)$. There is a simple way to check if a differential is exact. If the below equality holds, the differential is exact.

$$dF = f(x, y) dx + g(x, y) dy ; \quad \frac{\partial f}{\partial y} = \frac{\partial g}{\partial x} \quad (78)$$

All inexact differentials have a non-unique 'integrating factor' $r(x, y)$ which, when multiplied by the inexact differential, produces an exact differential. A simple way of finding the exact differential is to take the exact differential below and subject it to the equality condition stated previously.

$$dG = r(x, y) dF \quad (79)$$

2.2.2 Further Derivations

We can derive equations for dW and dQ

$$dW = -P dV ; \quad dQ = T dS \quad (80)$$

We can also derive the equation for dU by replacing E with the internal energy U in the differential form of the fundamental relation

$$dU = T dS - P dV + \mu dN \quad (81)$$

2.3 Thermodynamic Processes

[Sw 116-122]

Due to the Second Law, the total entropy must be maximized as a system tends towards equilibrium $dS \geq 0$. An 'irreversible' process is when the total entropy of a system increases. Running the process backwards is impossible since the total entropy cannot decrease.

We will call a 'quasi-static' process reversible. A quasi-static process is an idealization in which a system changes through a series of infinitesimal steps.

2.3.1 Heat Engines

We are interested in a heat engine in which a certain amount of heat dQ turns into work dW . Due to the First Law, $dW \leq dQ$.

We'll assume that the most-efficient engine has a quasi-static process. We'll require that heat is transferred from a reservoir at a high temperature T_H to a reservoir at low temperature $T_L < T_H$. The net work done is given by the heat exchanged

$$dW = dQ_H + dQ_L \quad (82)$$

The total entropy change can be described from the fundamental relation. For a quasi-static process, this is equal to 0.

$$dS = \frac{dQ_H}{T_H} + \frac{dQ_L}{T_L} = 0 \quad (83)$$

Combining the last two equations,

$$dW = \left(1 - \frac{T_L}{T_H}\right) dQ_H \quad (84)$$

We'll define the efficiency of a heat engine, η . The formula below is the Carnot efficiency. No heat engine, real or ideal, can have an efficiency higher than the Carnot efficiency.

$$\eta = \frac{dW}{dQ_H} = 1 - \frac{T_L}{T_H} \quad (85)$$

2.3.2 Refrigerators and Heat Pumps

Refrigerators are reversed heat engines. They remove heat from the reservoir at low temperature T_L and move it to that at high temperature T_H with the application of energy. The goal of a refrigerator is to remove as much heat from the low temperature reservoir. A heat pump works a similar way, but with the goal of maximizing the heat put into the high temperature reservoir. Their efficiencies are as follows

$$\epsilon_R = \frac{dQ_L}{-dW} = \frac{T_L}{T_H - T_L} ; \quad \epsilon_{HP} = \frac{-dQ_H}{-dW} = \frac{T_H}{T_H - T_L} \quad (86)$$

2.4 Thermodynamic Potentials

[Sw 123-132]

2.4.1 The Legendre Transform

A function f is called convex over a region $[a, b]$ if, for every choice of numbers $\{x_1, x_2\} \in [a, b]$ and every $t \in [0, 1]$

$$f(tx_1 + (1-t)x_2) \leq ty(x_1) + (1-t)y(x_2) \quad (87)$$

This effectively states that a straight line drawn between two points a and b on the function will lie above the function evaluated at all points in the range $[a, b]$. Locally, this suggests that convex functions must

have a positive second derivative $f''(x) \geq 0$. A concave function holds the analogous opposite of these conditions (flip the inequalities). A Legendre transform is therefore defined

$$p(x) = \frac{\partial y(x)}{\partial x} ; \quad y^*(p) \equiv \begin{cases} \min_x [y(x) - xp] & \text{if } y \text{ is convex} \\ \max_x [y(x) - xp] & \text{if } y \text{ is concave} \end{cases} \quad (88)$$

The Legendre transformation $y^*(p)$ thereby maps a concave or convex function $y(x)$ from its original input space x to an input space $\frac{\partial y}{\partial x}$. If a Legendre transform is performed on a region of $y(x)$ which is not totally convex or concave, the resultant transformation $y^*(p)$ loses some information from the original function. Some useful properties of the Legendre transform are

$$\frac{\partial y^*(p)}{\partial p} = -x(p) ; \quad y^{**} = y \quad (89)$$

We can go between an infinitesimal and its Legendre transform simply by switching the parameters and changing the sign.

$$dy = p dx ; \quad dq = -x dp \quad (90)$$

2.4.2 New Thermodynamic Potentials F, H and G

We can rearrange the fundamental relation to produce new thermodynamic potentials using the Legendre transform. This method swaps independent variables for dependent variables and will allow us the freedom to study systems when any arbitrary variables are held constant.

The Helmholtz free energy F is the Legendre transform of U with respect to temperature T .

The enthalpy H is the Legendre transform of U with respect to pressure P .

The Gibbs free energy G is the Legendre transform of U with respect to temperature T and pressure P . It can also be found by taking the Legendre transform of F with respect to P or H with respect to T .

All the thermodynamic potentials F , H , and G can be found by taking the fundamental relation for dU and performing some sort of Legendre transform. This is simple result is just exchanging a term's independent and dependent variable and changing the sign. A host of thermodynamic potentials, besides U , F , H , and G can be found this way. The below table summarizes the main representations of the fundamental relation.

Potential	Parameters	Differential Form
Energy	$U = U(S, V, N)$	$dU = TdS - PdV + \mu dN$
Helmholtz free energy	$F = F(T, V, N)$	$dF = -SdT - PdV + \mu dN$
Enthalpy	$H = H(S, P, N)$	$dH = TdS + VdP + \mu dN$
Gibbs free energy	$G = G(T, P, N)$	$dG = -SdT + VdP + \mu dN$

2.5 The Consequences of Extensivity

[Sw 133-137]

The entropy is extensive if

$$S(\lambda U, \lambda V, \lambda N) = \lambda S(U, V, N) \quad (91)$$

for any $\lambda > 0$. For an extensive system, S , U , V , and N are all proportional to the size of the system.

From this statement, we can derive the Euler equation, relevant for extensive systems.

$$U = TS - PV + \mu N \quad (92)$$

Taking the differential of the Euler equation and reducing it with the fundamental relation, we arrive at the Gibbs-Duhem relation.

$$0 = SdT - VdP + \sum_{j=1}^r N_j d\mu_j \quad (93)$$

This relation states that the three intensive parameters T , P , and μ are not independent for extensive systems. Note that the above form is the general Gibbs-Duhem relation for r components.

Lastly, we can see consequences of the Euler equation in the other thermodynamic potentials F , H , and G . For extensive systems,

Potential	If Extensive	
Energy	$U = TS - PV + \mu N$	
Helmholtz free energy	$F = U - TS$	$= -PV + \mu N$
Enthalpy	$H = U + PV$	$= TS + \mu N$
Gibbs free energy	$G = U - TS + PV$	$= \mu N$

2.6 Thermodynamic Identities

[Sw 138-155]

A quick note about notation: The statement $(\frac{\partial A}{\partial X})_Y$ is to say the partial derivative of A with respect to X while Y is held constant.

2.6.1 First and Second Derivatives

We can use the differentials of the thermodynamic potentials U , F , H , and G to show how parameters relate via partial derivatives. For example, using the fundamental relation, we can state the following first derivatives:

$$\left(\frac{\partial U}{\partial S}\right)_{V,N} = T ; \quad \left(\frac{\partial U}{\partial V}\right)_{S,N} = -P ; \quad \left(\frac{\partial U}{\partial N}\right)_{S,V} = \mu \quad (94)$$

We can construct a number of second derivatives of the above equations, but there will only be several independent ones per differential. This is because the order that we take the partial derivatives doesn't matter, i.e. $\frac{\partial^2 A}{\partial X \partial Y} = \frac{\partial^2 A}{\partial Y \partial X}$.

Typically, the composition of the system is fixed (N is constant). In this case, the standard second derivatives are given in Table 1.

Name	Symbol	Definition
Coefficient of thermal expansion	α	$\alpha = \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_{P,N}$
Isothermal compressibility	κ_T	$\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_{T,N}$
Specific heat per particle at constant pressure	c_P	$c_P = \frac{T}{N} \left(\frac{\partial S}{\partial T} \right)_{P,N}$
Specific heat per particle at constant volume	c_V	$c_V = \frac{T}{N} \left(\frac{\partial S}{\partial T} \right)_{V,N}$

Table 1: Standard set of second derivatives

2.6.2 Maxwell Relations

Maxwell relations are formed by applying the condition for exact differentials to the differentials of thermodynamic potentials. Since each thermodynamic potential has an exact differential and since we can manipulate the fundamental relation easily through Legendre transforms, this method proves very useful in relating partial derivatives of our system parameters.

For example, the condition for an exact differential applied to the fundamental relation at constant N produces the following Maxwell relation.

$$dU = TdS - PdV \quad \rightarrow \quad \left(\frac{\partial T}{\partial V} \right)_{S,N} = - \left(\frac{\partial P}{\partial S} \right)_{V,N} \quad (95)$$

This procedure can be followed for other Legendre transforms of the fundamental relation.

2.6.3 Manipulating Partial Derivatives

Unfortunately, the Maxwell relations are insufficient in relating the partial derivatives of all thermodynamic parameters. We can manipulate thermodynamic partial derivatives through the use of Jacobians to create more complex relationships.

The Jacobian will help us manipulate partial derivatives. Jacobians are defined as the determinant of a matrix of derivatives.

$$\frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \frac{\partial u}{\partial x} \frac{\partial v}{\partial y} - \frac{\partial u}{\partial y} \frac{\partial v}{\partial x} \quad (96)$$

The definition extends in an intuitive manner to Jacobians of higher parameters. From the definition of the Jacobian, we can see the following important properties.

$$\frac{\partial(u, v)}{\partial(x, y)} = -\frac{\partial(v, u)}{\partial(x, y)}; \quad \frac{\partial(u, y)}{\partial(x, y)} = \left(\frac{\partial u}{\partial x} \right)_y \quad (97)$$

The Jacobian is antisymmetric under an exchange of parameters. Also, when only one pair of parameters differs between the numerator and denominator, the Jacobian is equal to the partial derivative of those parameters with the rest taken to be constant.

We also can play with Jacobians like algebraic fractions.

$$\begin{aligned} \frac{\partial(u, v)}{\partial(x, y)} &= \frac{\partial(u, v)}{\partial(r, s)} \frac{\partial(r, s)}{\partial(x, y)}; & \frac{\partial(u, v)}{\partial(x, y)} \frac{\partial(a, b)}{\partial(c, d)} &= \frac{\partial(u, v)}{\partial(c, d)} \frac{\partial(a, b)}{\partial(x, y)} \\ \frac{\partial(u, v)}{\partial(r, s)} &= 1 \bigg/ \frac{\partial(r, s)}{\partial(u, v)}; & \left(\frac{\partial u}{\partial x} \right)_y &= 1 \bigg/ \left(\frac{\partial x}{\partial u} \right)_y \end{aligned} \quad (98)$$

The general strategy for deriving thermodynamic identities is to use the above rules to manipulate an unknown partial derivative into derivatives which you have information for. This typically involves introduction of the standard set of second derivatives α , κ_T , c_P , or c_V .

Sw gives a general strategy to reduce a given partial derivative to an algebraic expression containing only the standard set of partial derivatives.

1. Express the partial derivative as a Jacobian
2. If there are any thermodynamic potentials, U , F , H , or G , bring them to the numerator. It is also useful to insert $\partial(T, P)$. This is because the standard set of partial derivatives are typically expressed as derivatives with respect to T or P , with the other held constant.
3. Eliminate thermodynamic potentials if you know the derivative.
4. If you don't know the derivative of the thermodynamic potential, use the fundamental relation to get replace it.

$$\left(\frac{\partial F}{\partial P} \right)_{T, N} = -S \left(\frac{\partial T}{\partial P} \right)_{T, N} - P \left(\frac{\partial V}{\partial P} \right)_{T, N} + \mu \left(\frac{\partial N}{\partial P} \right)_{T, N} = -P \left(\frac{\partial V}{\partial P} \right)_{T, N} \quad (99)$$

5. If the system is extensive, bring μ to the numerator and eliminate it using the Gibbs-Duhem relation.
6. Move the entropy to the numerator and eliminate it either with specific heat (if the derivative is with respect to T) or using a Maxwell relation (if the derivative is with respect to P).
7. Move V to the numerator and eliminate the partial derivative in favor of α or κ_T .
8. Unless specifically asked, eliminate c_V in favor of c_P .

2.6.4 Some Identity Derivations

Sw does some examples of this procedure in [Sw 148-151]. The examples are useful to look at and practice, but largely resemble the procedure described above.

Sw first describes how the temperature of a gas changes undergoing a change in pressure at constant enthalpy. The partial derivative has a name: the Joule-Thomson coefficient μ_{JT} .

$$\mu_{JT} = \left(\frac{\partial T}{\partial P} \right)_{H,N} = \frac{V}{Nc_P}(T\alpha - 1) \quad (100)$$

Sw also shows that c_P and c_V are not independent, and are instead related by the following.

$$c_P = c_V + \frac{\alpha^2 TV}{N\kappa_T} \quad (101)$$

2.7 Extremum Principles and Stability Conditions

[Sw 156-166]

We now have all the tools necessary to study how a system reaches equilibrium under various conditions.

2.7.1 Extremum Principles

When we introduced entropy, we showed that a thermally-isolated system with $dU = 0$ moves to maximize entropy S . We can derive extremum principles for other system conditions.

- When energy is constant $dU = 0$ and the system is thermally-isolated $dQ = 0$, entropy S is maximized.
- When entropy is constant $dS = 0$ and the system is thermally-isolated $dQ = 0$, energy U is minimized.
- If the temperature is held constant $dT = 0$, the Helmholtz free energy F is minimized.
- If the pressure $dP = 0$ is held constant, the enthalpy H is minimized.
- If the temperature and pressure are held constant $dT = dP = 0$, the Gibbs free energy G is minimized.

2.7.2 Stability Conditions

We can use the extremum principles to show how certain thermodynamic parameters vary when moving towards equilibrium. Sw gives derivations for each of the following cases in Sw 168-175. The derivations are useful to familiarize yourself with.

In a free, equilibrium-seeking system, the following inequalities for second derivatives hold.

$$\begin{aligned} \kappa_S = -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_{S,N} &> 0 ; & \kappa_T &> 0 \\ c_V &> 0 ; & c_P &> 0 \end{aligned} \quad (102)$$

Using the second derivative inequalities, we can put constraints on second derivatives of thermodynamic potentials valid at stability.

$$\begin{aligned} U(S, V, N) : & \quad \left(\frac{\partial^2 U}{\partial S^2} \right)_{V,N} > 0 ; & \left(\frac{\partial^2 U}{\partial V^2} \right)_{S,N} &> 0 \\ F(T, V, N) : & \quad \left(\frac{\partial^2 F}{\partial T^2} \right)_{V,N} < 0 ; & \left(\frac{\partial^2 F}{\partial V^2} \right)_{T,N} &> 0 \\ H(S, P, N) : & \quad \left(\frac{\partial^2 H}{\partial S^2} \right)_{P,N} > 0 ; & \left(\frac{\partial^2 H}{\partial P^2} \right)_{S,N} &< 0 \\ G(T, P, N) : & \quad \left(\frac{\partial^2 G}{\partial T^2} \right)_{P,N} < 0 ; & \left(\frac{\partial^2 G}{\partial P^2} \right)_{T,N} &< 0 \end{aligned} \quad (103)$$

2.8 Phase Transitions

[Sw 177-193]

Phase transitions occur when the entropy undergoes a discontinuous change upon a change in extensive parameters. First-order phase transitions occur when the extensive variables change discontinuously (e.g. water freezing). Higher-order phase transitions occur when higher order partial derivatives of the extensive variables change discontinuously (e.g. magnetization approaching a critical temperature). For this section, we will focus on first-order phase transitions.

2.8.1 The van der Waals Fluid

The van der Waals model of a real gas suggests an interaction potential between particles. The van der Waals model has both attractive and repulsive parts. The repulsive force dominates on short distances while the attractive force is more prominent at long distances. We will assume that the strength of the attractive and repulsive interactions varies with the parameters a and b , respectively.

We will start with the ideal gas and make the following further approximations. We expect that the average energy of attraction to be proportional to the density of particles in the neighbor of any given particle, which we will approximate by the average density N/V . We will then assume that the most important effect of the repulsive interaction is a constant reduction in the available volume. Thus, we make the following substitutions

$$U \rightarrow U - a \left(\frac{N^2}{V} \right) ; \quad V \rightarrow V - bN \quad (104)$$

Substituting these into the Helmholtz free energy for the ideal gas, we arrive at the free energy for the van der Waals fluid.

$$F_{vdW} = -Nk_B T \left[\ln \left(\frac{V - bN}{N} \right) + \frac{3}{2} \ln(k_B T) + X \right] - a \left(\frac{N^2}{V} \right) \quad (105)$$

Taking the partial derivatives of F_{vdW} with respect to T and V , we arrive at two equations of state.

$$P = \frac{Nk_B T}{V - bN} - \frac{aN^2}{V^2} ; \quad U = \frac{3}{2}Nk_B T - a \left(\frac{N^2}{V} \right) \quad (106)$$

Note that the van der Waals fluid is an extensive system.

At high temperatures, a plot of $P(V)$ is monotonically decreasing with negative slope. However, below an inflection point at T_c , the relationship looks like Figure 1. The inflection point at which this occurs is given by the following critical values.

$$V_c = 3bN ; \quad P_c = \frac{a}{27b^2} ; \quad k_B T_c = \frac{8a}{27b} \quad (107)$$

We can rewrite the equations of state in a dimensionless form using these values: $\tilde{V} = V/V_c$, $\tilde{P} = P/P_c$, and $\tilde{T} = T/T_c$.

$$\tilde{P} = \frac{8\tilde{T}}{3\tilde{V} - 1} - \frac{3}{\tilde{V}^2} ; \quad (\tilde{P} + 3\tilde{V}^{-2})(3\tilde{V} - 1) = 8\tilde{T} \quad (108)$$

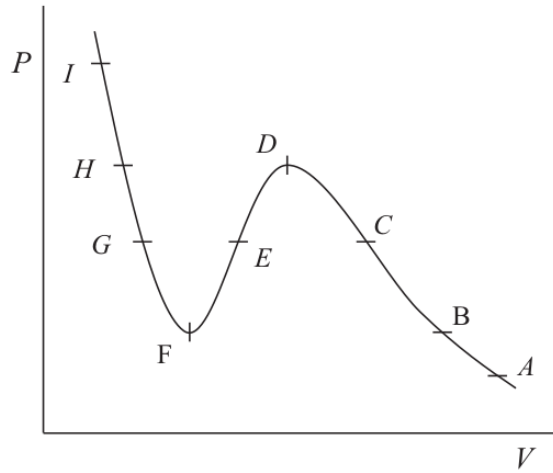


Figure 1: Schematic $P(V)$ plot of an isotherm for the van der Waals equation at a low temperature.

2.8.2 Instabilities and the Liquid-Gas Transition

It is clear that the region between points B and H is problematic. Volume is a triple-valued function of the pressure in this region. Which is correct?

For the stable state, the Gibbs free energy should be a minimum. We can use the fact that the system is

extensive to show the following relationship.

$$G = \mu N = \int V dP \quad (109)$$

Following $V(P)$ as in the inverse of Figure 1, we arrive at $G(P)$, shown in Figure 2.

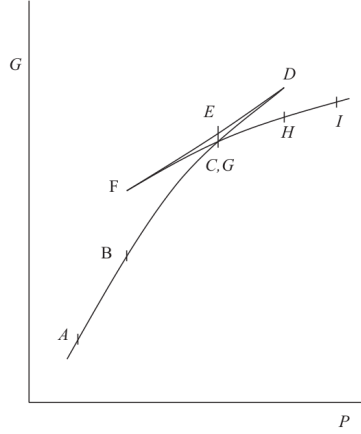


Figure 2: Schematic $G(P)$ plot of an isotherm for a van der Waals fluid at a low temperature.

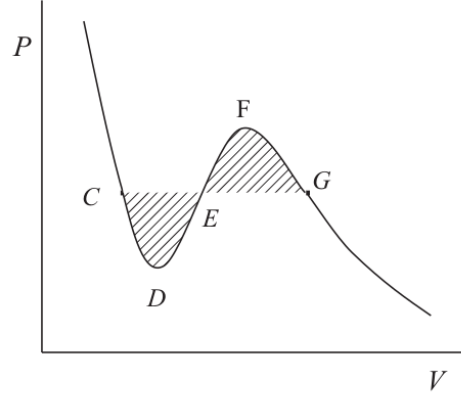


Figure 3: Schematic $P(V)$ plot of an isotherm for the van der Waals equation at a low temperature.

We see in Figure 2 that there exists a high G loop around the transition pressure. Since G should be minimized, passing through the transition pressure should move the system directly from point C to G . The corresponding $P(V)$ diagram replaces the curve between C and G with a flat horizontal line. This horizontal line is at an equilibrium P where the areas of the two shaded regions are equal. The procedure of adjusting the position of the line until the two areas are equal is known as the Maxwell construction.

Below point C , the system is in a single, distinct, high-density phase. We call this the liquid phase. Above point G , the system is in a single, low-density phase, called the gas phase. In between C and G , the system is in partial states of liquid and gas.

The van der Waals equations can be linked to produce an expression for $P(T)$ at constant V . This relationship produces a phase diagram shown in Figure 4. At temperatures or pressures above the critical point, there is no distinction between liquid and gas; there is only a fluid. Below T_c and P_c , the system must undergo a rapid, discontinuous change to move between liquid and gas. The liquid-gas phase transition requires adding energy. The latent heat L necessary to move from point C (liquid) to point G (gas) in figure 3 is given by the following.

$$L = T\Delta S = T \int_C^G \left(\frac{\partial P}{\partial T} \right)_{V,N} dV \quad (110)$$

The partial derivative can be calculated from the van der Waals equation of state.

2.9 The Nernst Postulate

[Sw 194-198]

The Nernst Postulate: **The entropy of a thermodynamic system goes to a constant as the temperature goes to zero.**

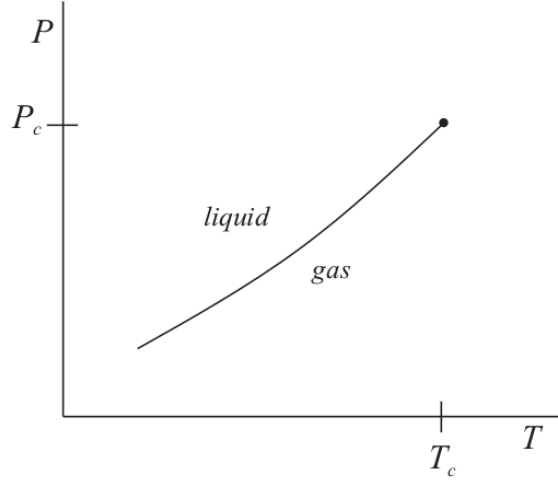


Figure 4: Schematic $P(T)$ plot of the phase diagram of the van der Waals fluid.

This statement is valid for all real systems. However, it fails for all classical systems (e.g. the ideal gas). The reason for the validity of the Nernst Postulate for real system lies entirely in quantum statistical mechanics.

However, in the meantime, we can derive the following consequences of the Nernst Postulate.

$$\lim_{T \rightarrow 0} c_V(T) = \lim_{T \rightarrow 0} c_P(T) = 0 ; \quad \lim_{T \rightarrow 0} \alpha(T) = 0 \quad (111)$$

3 Classical Statistical Mechanics

3.1 Ensembles in Classical Statistical Mechanics

[Sw 201-226]

We previously described the classical microcanonical ensemble by a uniform probability in phase space, subject to the constraints that the particles are all in a particular volume and that the total energy is constant.

The entropy of a classical system of interacting particles is given by an integral over all phase space.

$$S(E, V, N) = k \ln \left[\frac{1}{h^{3N} N!} \int dq \int dp \delta(E - H(p, q)) \right] \quad (112)$$

The integrals $\int dq \int dp$ are over the $6N$ -dimensional phase space, which makes them difficult or impossible to carry out explicitly for interacting systems. We therefore will look to further statistical methods of describing classical systems.

3.1.1 Canonical Ensemble

The canonical ensemble is the probability distribution in phase space for a system in contact with a thermal reservoir at a known temperature. The difference between the microcanonical case is that we will calculate

a probability distribution over states in our system, instead of assuming a uniform probability density over phase space.

First, we'll calculate the energy distribution for the system of interest. Assume that the system of interest is in thermal contact with a reservoir at temperature T and at equilibrium. The probability distribution is given by

$$P(E) = \frac{1}{Z} \Omega(E) \exp\{-\beta E\} ; \quad Z(T, V, N) = \int \Omega(E, V, N) \exp\{-\beta E\} dE ; \quad \beta = \frac{1}{k_B T} \quad (113)$$

where $\Omega(E)$ is the multiplicity of the system when it has energy E . The function $Z(T, V, N)$ is called the 'partition function,' and it is simply a normalization constant for the probability distribution. β is a common convention for expressing the temperature dependence that we will use in the future.

We can derive the same canonical ensemble probability density over points in phase space. Denote the set of momentum and position variables describing the system of interest as $\{p, q\}$. Assign a Hamiltonian $H(p, q)$ to the energy of the system of interest. Using the general equation for $\Omega(E, V, N)$, we arrive at

$$P(p, q) = \frac{1}{h^{3N} N! Z} \exp\{-\beta H(p, q)\} ; \quad Z = \frac{1}{h^{3N} N!} \int dq \int dp \exp\{-\beta H(p, q)\} \quad (114)$$

As you can see, the canonical phase space probability distribution is the same as the energy probability distribution, except with $E \rightarrow H(p, q)$ and an extra factor of $(h^{3N} N!)^{-1}$ in the partition function to account for the volume of phase space. While similar, the phase space probability distribution will be easier to work with in many cases.

3.1.2 Liouville Theorem and Consequences of the Canonical Ensemble

The Liouville Theorem: **The canonical probability distribution does not change with time (if the system remains at equilibrium).**

Sw provides a rigorous proof of the Liouville theorem in [Sw 207-209], but a simple explanation is relies on the fact that the canonical probability density depends only on the total energy. Since the trajectory of a point in phase space conserves energy, the canonical probability distribution does not change with time.

Sw derives the width of the probability distribution for the energy $P(E)$.

$$\Omega(E) \propto E^f ; \quad \sigma_E = \frac{E_{eq}}{\sqrt{f}} \quad (115)$$

where E_{eq} is the equilibrium energy of the system. Typically, the value of f is approximately proportional to N (for the ideal gas, $f = 3N/2$). This means that $\sigma_E \propto N^{-1/2}$. This important result means that the width of the probability distribution is very small for a many particle system.

There are several very useful thermodynamic identities related to the partition function Z .

$$F = -k_B T \ln Z ; \quad U = \langle E \rangle = -\frac{\partial}{\partial \beta} \ln Z = \frac{\partial(\beta F)}{\partial \beta} \quad (116)$$

The partition function of the classical ideal gas where $[\phi(\mathbf{r}_i, \mathbf{r}_j = 0)]$ is given by

$$Z = \frac{1}{h^{3N} N!} (2\pi m k_B T)^{3N/2} V^N \quad (117)$$

3.1.3 Factorization of the Partition Function

A useful advantage of the canonical partition function is that the integrals over the momenta can be carried out exactly for any system in which the forces do not depend on the momenta. If $H(p, q)$ only depends on the momenta through the kinetic energy, the partition function becomes

$$\begin{aligned}
Z &= \frac{1}{h^{3N} N!} \int dp \int dq \exp \left\{ -\beta \left(\sum_{j=1}^N \frac{|\mathbf{p}_j|^2}{2m} + \sum_{j=1}^N \sum_{i>j}^N \phi(\mathbf{r}_i, \mathbf{r}_j) \right) \right\} \\
&= \frac{1}{h^{3N} N!} \cdot \int dp \exp \left\{ -\beta \sum_{j=1}^N \frac{|\mathbf{p}_j|^2}{2m} \right\} \cdot \int dq \exp \left\{ -\beta \sum_{j=1}^N \sum_{i>j}^N \phi(\mathbf{r}_i, \mathbf{r}_j) \right\} \\
&= \frac{1}{h^{3N} N!} (2\pi m k_B T)^{3N/2} \cdot \int dq \exp \left\{ -\beta \sum_{j=1}^N \sum_{i>j}^N \phi(\mathbf{r}_i, \mathbf{r}_j) \right\}
\end{aligned} \tag{118}$$

Furthermore, the expression for the partition function can be greatly simplified if the Hamiltonian can be written as the sum of N independent parts. This is particularly relevant in the case of non-interacting particles.

$$\begin{aligned}
Z &= \frac{1}{h^{3N} N!} \int dp \int dq \exp \left\{ -\beta \left(\sum_{j=1}^N \frac{|\mathbf{p}_j|^2}{2m} + \sum_{j=1}^N \phi(\mathbf{r}_j) \right) \right\} \\
&= \frac{1}{h^{3N} N!} (2\pi m k_B T)^{3N/2} \cdot \int dq \exp \left\{ -\beta \sum_{j=1}^N \phi(\mathbf{r}_j) \right\} \\
&= \frac{1}{h^{3N} N!} (2\pi m k_B T)^{3N/2} \prod_{j=1}^N \int d\mathbf{r}_j \exp \{ -\beta \phi_j(\mathbf{r}_j) \}
\end{aligned} \tag{119}$$

If all the $\int d\mathbf{r}_j$ integrals are the same (e.g. the particles are identical), the result simplifies to

$$Z = \frac{1}{h^{3N} N!} (2\pi m k_B T)^{3N/2} \left(\int d\mathbf{r}_1 \exp \{ -\beta \phi_1(\mathbf{r}_1) \} \right)^N \tag{120}$$

3.1.4 Simple Harmonic Oscillator

The partition function for a single, one-dimensional, simple harmonic oscillator (SHO) with hamiltonian H_1 given. This is a straightforward calculation, as the expression only contains two Gaussian integrals.

$$H_1 = \frac{1}{2} K x^2 + \frac{p^2}{2m} ; \quad Z_1 = \frac{1}{\beta \hbar \omega} ; \quad \omega = \sqrt{\frac{K}{m}} \tag{121}$$

We can generalize this to N independent, SHO's. This is easy because the Hamiltonian is separable.

$$H_N = \sum_{j=1}^N \left(\frac{1}{2} K_j x_j^2 + \frac{p_j^2}{2m_j} \right) ; \quad Z_N = \prod_{j=1}^N \left(\frac{1}{\beta \hbar \omega_j} \right) ; \quad \omega_j = \sqrt{\frac{K_j}{m_j}} \tag{122}$$

Obviously, if all N SHO's have the same frequency, $Z_N = (\beta \hbar \omega_j)^{-N}$.

3.2 Grand Canonical Ensemble

[Sw 227-233]

The grand canonical ensemble describes a system that can exchange both energy and particles with a reservoir. We'll assume that the reservoir and the system of interest can exchange both energy and particles with each other, but are completely isolated from the rest of the universe. The grand canonical probability distribution $P(E, N)$ and partition function \mathcal{Z} is given below.

$$P(E, N) = \frac{1}{\mathcal{Z}} \Omega(E, N) \exp\{-\beta E + \beta \mu N\} ; \quad \mathcal{Z} = \sum_{N=0}^{\infty} \int_0^{\infty} dE \Omega(E, N) \exp\{-\beta E + \beta \mu N\} \quad (123)$$

where $\Omega(E, N)$ is the multiplicity of the system at a given E and N .

We can also express the grand canonical partition function as a Laplace transform of the canonical partition function.

$$\mathcal{Z}(T, V, \mu) = \sum_{N=0}^{\infty} Z(T, V, N) \exp\{\beta \mu N\} \quad (124)$$

where $Z(T, V, N)$ is the canonical partition function for a given N .

The grand canonical partition function is related to the Legendre transform of U with respect to T and μ .

Furthermore, for an extensive system, it is related to PV .

$$U[T, \mu] = -k_B T \ln \mathcal{Z} ; \quad k_B T \ln \mathcal{Z} = PV \quad (125)$$

The grand canonical partition function for an ideal gas is given by

$$\mathcal{Z} = \exp\left\{(2\pi m k_B T)^{3/2} h^{-3} V e^{\beta \mu}\right\} \quad (126)$$

3.3 Irreversibility

[Sw 234-244]

Sw presents a historical account of the debate on irreversibility proof as well as a modern interpretation. He discusses the free expansion of the classical ideal gas and gives an approximate proof of Liouville's Theorem. Much of this information is not applicable in this course. See [Sw 234-243].

4 Quantum Statistical Mechanics

4.1 Quantum Ensembles

[Sw 247-281]

The key difference between classical and quantum systems is in their microscopic states. While classical states determine exact position and momentum for each particle, quantum states can only provide probabilities for observables. This affects the calculation of probabilities which derive thermodynamic parameters.

4.1.1 Basic Quantum Mechanics

The microscopic state of a quantum system is described by a wave function, rather than a point in phase space. For a one-particle system, the wave function is a complex function $\psi(\mathbf{r}, t)$. The probability density for the position measurement of the particle is given by $|\psi(\mathbf{r}, t)|^2$.

We will use a simplified bracket notation from here on. ψ^* denotes the complex conjugate of the function ψ . We introduce the 'bra', 'ket', and their combination respectively.

$$\langle\psi| = \psi^*(\mathbf{r}, t) ; \quad |\psi\rangle = \psi(\mathbf{r}, t) ; \quad \langle\psi_A| X |\psi_B\rangle = \int \psi_A^*(\mathbf{r}, t) X \psi_B(\mathbf{r}, t) d^3r \quad (127)$$

where X is some operator that acts on ψ_B . The momentum, \mathbf{p} , is represented by a vector operator rather than a real vector.

$$\mathbf{p} = -i\hbar\nabla = -i\hbar\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) ; \quad \langle\mathbf{p}\rangle = \langle\psi|\mathbf{p}|\psi\rangle = -i\hbar \int \psi^* \nabla \psi d^3r \quad (128)$$

The kinetic and potential energy are represented by a Hamiltonian $H(\mathbf{r}, \mathbf{p})$.

$$H = H(\mathbf{r}, \mathbf{p}) = -\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) ; \quad \nabla^2 = \nabla \cdot \nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad (129)$$

The wave function must satisfy the Schrodinger equation

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) = H \psi(\mathbf{r}, t) \quad (130)$$

Since the Hamiltonian is time-independent, we can write it in a form that separates the functions of time and space. Below is the time-independent Schrodinger equation.

$$\psi(\mathbf{r}, t) = \psi(\mathbf{r}) e^{-iEt/\hbar} ; \quad H \psi(\mathbf{r}) = E \psi(\mathbf{r}) \quad (131)$$

This means that $\langle H \rangle = E$. Usually, the time-independent Schrodinger equation only has solutions for certain values of E . We typically denote these by a set of d quantum numbers for a d -dimensional system. For the one dimensional case, we will use n as a quantum number. For specific eigenvalues E_n , we have eigenfunctions ψ_n .

$$|n\rangle = \psi_n(\mathbf{r}) e^{-iE_n t/\hbar} ; \quad H |n\rangle = E_n |n\rangle ; \quad \langle n|m\rangle = \delta_{n,m} \quad (132)$$

The last identity above states that the set of $\{|n\rangle\}$ is orthonormal. We can express any wave function as an expansion in the set of all eigenfunctions.

$$|\psi\rangle = \sum_n |n\rangle c_n ; \quad c_n = \langle n|\psi\rangle \quad (133)$$

$|c_n|^2$ is the probability that the system would be found in the state $|n\rangle$ after a measurement has been made that put the system into an eigenstate. We can write the expectation value as the energy as $\langle H \rangle = \sum_n E_n |c_n|^2$.

Our final expression for an arbitrary wave function can be written in terms of the expansion in eigenfunctions.

$$|\psi, t\rangle = \sum_n c_n e^{-iE_n t/\hbar} |n\rangle \quad (134)$$

Quantum systems that contain many particles are also described by a single wave function, but one that is a function of the coordinates of every particle in the system, $\psi(\{\mathbf{r}_j | j = 1, \dots, N\})$.

4.1.2 Model Probabilities and Density Matrix

We now make a distinction between the probability distribution of a wave function and the probability distribution of each quantum state. The latter we will call P_ψ . It refers to the probability that the system will be in state ψ . The expectation value of any operator \mathcal{A} in a many-body system can now be written as

$$\langle \mathcal{A} \rangle = \int_{\psi} P_{\psi} \langle \psi | \mathcal{A} | \psi \rangle = \sum_n P_n \langle n | \mathcal{A} | n \rangle \quad (135)$$

in the last identity above, we have split up P_ψ into a probability distribution over orthonormal eigenstates P_n .

A common representation of averages in a quantum ensemble is provided by the density matrix.

$$\rho = \int_{\psi} P_{\psi} |\psi\rangle \langle \psi| ; \quad \rho_{n,m} = \int_{\psi} P_{\psi} \langle m | \psi \rangle \langle \psi | n \rangle \quad (136)$$

where $\rho_{n,m}$ is the (n, m) -th element in the matrix representation of ρ . The expectation value of \mathcal{A} is then given simply by

$$\langle \mathcal{A} \rangle = \text{Tr}\{\rho \mathcal{A}\} \quad (137)$$

Although the density matrix contains enough information to calculate any ensemble average, it does not uniquely specify the quantum ensemble.

4.1.3 Quantum Canonical Ensemble

Like its classical counterpart, the quantum canonical ensemble describes the behavior of a system in contact with a thermal reservoir. The energy probability density and partition function are given below

$$P_n = \frac{1}{Z} \exp\{-\beta E_n\} ; \quad Z = \sum_n \exp\{-\beta E_n\} \quad (138)$$

This is very similar to the classical canonical partition function, but the energy levels are discrete. If the energy states are degenerate with degeneracy $\Omega(l)$, the partition function is given by

$$Z = \sum_l \Omega(l) \exp\{-\beta E_l\} \quad (139)$$

4.1.4 Consequences of the Quantum Canonical Ensemble

We see the same thermodynamic identities as in the classical, continuous case. We also see an alternative expression for entropy S .

$$U = - \left(\frac{\partial}{\partial \beta} \ln Z \right)_{V,N} ; \quad F = -k_B T \ln Z ; \quad S = -k_B \sum_n P_n \ln P_n \quad (140)$$

Using this last identity, Sw proves the Liouville Theorem in [Sw 264-265].

A useful mathematical implication of the quantum canonical distribution is that the temperature derivative can be expressed as

$$\frac{\partial}{\partial \beta} \langle \mathcal{A} \rangle = -[\langle \mathcal{A} E \rangle - \langle \mathcal{A} \rangle \langle E \rangle] \quad (141)$$

This general result works for any operator \mathcal{A} . Using this, we can show

$$\frac{\partial}{\partial \beta} \langle H \rangle = -\sigma_E^2 ; \quad c_V = \frac{1}{N k_B T^2} \sigma_E^2 \quad (142)$$

4.1.5 Factorization of the Partition Function

Just like as in the classical case, it will prove very useful to factorize the Hamiltonian into independent, low-dimensional parts.

$$H = \sum_{j=1}^N H_j ; \quad H_j |n_j\rangle = E_{n_j} |n_j\rangle \quad (143)$$

This will allow us to write the partition function in the following simple form

$$Z = \sum_{\{n_j\}} \prod_{j=1}^N \exp\{-\beta E_{n_j}\} = \prod_{j=1}^N \left(\sum_{n_j} \exp\{-\beta E_{n_j}\} \right) \quad (144)$$

The exchange of the \sum and \prod is allowable in this case (N is finite). Note $\sum_{\{n_j\}}$ is the sum over the set of all quantum numbers for the entire system, while \sum_{n_j} is the sum over the quantum numbers associated with the term H_j .

4.1.6 Special Systems

There are two special systems which are specifically important to study under the quantum canonical distribution.

The first is a two-level energy system. Consider a Hamiltonian $H = \epsilon n$ where $n \in \{0, 1\}$. The partition function, average number of excitations, free energy, and energy are given below

$$\begin{aligned} Z &= 1 + \exp\{-\beta \epsilon\} ; & \langle n \rangle &= \frac{1}{\exp\{\beta \epsilon\} + 1} \\ F &= -k_B T \ln(1 + \exp\{-\beta \epsilon\}) ; & U &= \frac{\epsilon}{\exp\{\beta \epsilon\} + 1} = \epsilon \langle n \rangle \end{aligned} \quad (145)$$

The second is a single, one-dimensional single harmonic oscillator SHO. The Hamiltonian, energy levels, and angular frequency are given below

$$H = \frac{1}{2} K x^2 - \frac{\hbar^2}{2m} \frac{d}{dx^2} ; \quad E_n = \hbar \omega \left(n + \frac{1}{2} \right) ; \quad \omega = \sqrt{\frac{K}{m}} \quad (146)$$

This produces the following partition function, average number of excitations, free energy, and energy

$$\begin{aligned} Z &= \frac{\exp\{-\beta\hbar\omega/2\}}{1 - \exp\{-\beta\hbar\omega\}} ; & \langle n \rangle &= \frac{1}{\langle \beta\hbar\omega \rangle - 1} \\ F &= \frac{1}{2}\hbar\omega + k_B T \ln(1 - \exp\{-\beta\hbar\omega\}) ; & U &= \frac{1}{2}\hbar\omega + \frac{\hbar\omega}{\exp\{\beta\hbar\omega\} - 1} = \frac{1}{2}\hbar\omega + \hbar\omega \langle n \rangle \end{aligned} \quad (147)$$

Sw relates these two general cases to specific examples. He explains magnetic energy due to a spin one-half particle and Einstein modes in a crystal. See [Sw 270-271] and [Sw 273-275].

4.2 Black-Body Radiation

[Sw 282-290]

'Black body' refers to an object that absorbs all radiation incident on it and reflects nothing. However, it radiates light arising from thermal energy. The purpose of this chapter is to calculate the spectrum of radiation emanating from a black body.

Due to the Second Law of Thermodynamics, the frequency spectrum for all black bodies must be the same.

4.2.1 A Simple Model

Consider a cubic cavity with dimensions $L \times L \times L$. The sides are made of metal and it contains electromagnetic radiation, but no matter. Radiation can only come in and out of the cavity through a very small hole in one side. The only thing inside the cavity is electromagnetic radiation at temperature T . From electrodynamics, the electric field $\mathbf{E}(\mathbf{r})$ is governed by the wave equation. The 3D wave equation and the solution for the x -component is given below.

$$\nabla^2 \mathbf{E}(\mathbf{r}, t) = \frac{1}{c^2} \frac{\partial^2 \mathbf{E}(\mathbf{r}, t)}{\partial t^2} ; \quad E_x(\mathbf{r}, t) = E_{x,0} \sin(\omega t) \cos(k_x x) \sin(k_y y) \sin(k_z z) \quad (148)$$

The boundary conditions for this system implies that the parallel components of the electric field vanish at the walls of the cube. This imposes strict conditions on the allowable values of the k 's and ω .

$$k_x L = n_x \pi ; \quad k_y L = n_y \pi ; \quad k_z L = n_z \pi ; \quad k_x^2 + k_y^2 + k_z^2 = \frac{\omega^2}{c^2} \quad (149)$$

where $n_x, n_y, n_z \in \mathbb{Z}$. The value of ω thus depends on the vector $\mathbf{n} = (n_x, n_y, n_z)$.

$$\omega_{\mathbf{n}} = \frac{\pi c}{L} \sqrt{n_x^2 + n_y^2 + n_z^2} = \frac{n\pi c}{L} = \omega_n \quad (150)$$

where $n = |\mathbf{n}|$.

4.2.2 Black Body Energy Spectrum

Since each photon with frequency ω has energy $\hbar\omega$, the expectation value of energy in a given ω energy state is identical to that of the simple harmonic oscillator.

$$\langle \epsilon_\omega \rangle = \frac{\hbar\omega}{\exp\{\beta\hbar\omega\} - 1} \quad (151)$$

We assume that the system is uniformly distributed in \mathbf{n} -space $P_{\mathbf{n}}(\mathbf{n}) = 1$. We can then use a variable transformation to find the density of states as a function of frequency.

$$P_{\omega}(\omega) = 2\frac{1}{8} \int_0^{\infty} 4\pi n^2 dn \delta(\omega - nc\pi/L) = \pi \left(\frac{L}{c\pi} \right)^3 \omega^2. \quad (152)$$

Note, we accounted for the two polarizations of EM waves by multiplying the above expression by 2. We can then find the energy density spectrum as the product of the density of states for ω and the expected energy of the ω states.

$$u_{\omega} = \langle \epsilon_{\omega} \rangle P_{\omega}(\omega) = \frac{\hbar}{\pi^2 c^3} \omega^3 (\exp\{\beta \hbar \omega\} - 1)^{-1} \quad (153)$$

Returning to the assumption that radiation escapes through a small hole in one side, the radiated power j_{ω} is proportional to the energy density u_{ω} . The proportional factor is $c/4$, related to geometric constraints.

This gives us the Planck law for black-body radiation.

$$j_{\omega} = \frac{1}{4} c u_{\omega} = \left(\frac{\hbar}{4\pi^2 c^2} \right) \frac{\omega^3}{\exp\{\beta \hbar \omega\} - 1} \quad (154)$$

The plot of this function is shown in Figure 5. This function peaks at $\omega_{max} \approx 2.82144 k_B T / \hbar$.

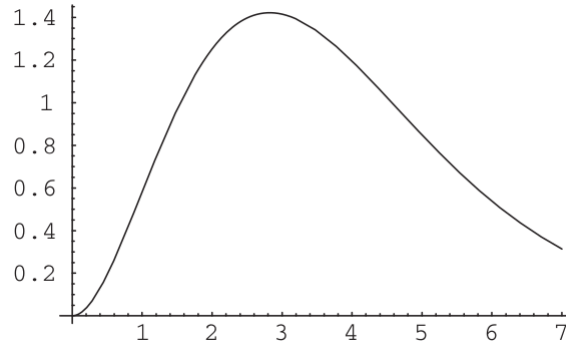


Figure 5: Plot of black-body radiation spectrum in dimensionless units. Note that $x = \beta \hbar \omega$. This function peaks at $x_{max} \approx 2.82144$.

For small values of $\omega \ll \omega_{max}$, the Planck law is quadratic in ω .

$$j_{\omega \ll \omega_{max}} \approx \left(\frac{1}{4\pi^2 c^2} \right) \omega^2 k_B T \quad (155)$$

4.2.3 Total Energy and Radiation

The average energy density over all ω can be found by integrating u_{ω} over all ω . We see that the total energy of a black body is proportional to T^4 .

$$\frac{U}{V} = u = \left(\frac{\pi^2}{15 \hbar^3 c^3} \right) (k_B T)^4 \quad (156)$$

The total black-body radiation can be found by integrating over j_ω or simply by multiplying u by the geometric factor $c/4$.

$$J_U = \sigma_B T^4 ; \quad \sigma_B = \frac{\pi^2 k_B^4}{60 \hbar^3 c^2} T^4 \quad (157)$$

where σ_B is the Stefan-Boltzmann constant.

4.3 The Harmonic Solid

[Sw 291-307]

A one-dimensional model of a crystal lattice is described by uniformly spaced points along a line R_j . Atoms are located at points r_j .

$$R_j = ja ; \quad r_j = R_j + x_j = ja + x_j \quad (158)$$

where a is the lattice constant, $j \in \mathbb{Z}$, and x_j is the deviation of the position of an atom relative to its associated lattice point. The energy of a microscopic state is given by

$$E = \frac{1}{2}m \sum_j \dot{x}_j^2 + \frac{1}{2}K \sum_j (x_j - x_{j+1})^2 \quad (159)$$

The energy is composed of a kinetic energy term and an interaction term. The energy is minimized when the atoms are equally spaced with $x_j = 0$.

4.3.1 Normal Modes

We will use the Fourier transform to reduce this problem into a separable system. The Fourier transform of the positions is given by

$$x_k = N^{-1/2} \sum_j x_j \exp\{-ikR_j\}; \quad k = \frac{2\pi}{\lambda} \quad (160)$$

where λ is the wavelength of the wave. These x_k are called the normal modes of the position variables. To convert from normal modes to positions, the inverse Fourier transform is given

$$x_j = N^{-1/2} \sum_k x_k \exp\{ikR_j\} \quad (161)$$

The following identity is very useful

$$\sum_j \exp\{i(k' - k)R_j\} = N\delta_{k,k'} \quad (162)$$

4.3.2 Boundary Conditions

The two boundary conditions we will consider are pinned and periodic. First, pinned boundary conditions suggest that the ends of the 1-D chain of atoms are pinned to the lattice spacing $x_0 = x_{N+1} = 0$. Our solutions then take the form

$$x_k = \sin(kR_j) ; \quad k = \frac{n\pi}{L} \quad (163)$$

where $n \in \mathbb{Z}$ and $L = Na$. Since we only have N atoms in the system, we only expect to find N independent solutions. In fact, the standard choice for pinned boundary conditions is to only use solutions which corresponding to values of n from 1 to N .

The other boundary conditions are periodic (e.g. a circular loop of atoms). In these boundary conditions, $R_{j+N} = R_j$. Our solutions are in the form

$$x_k = \exp\{ikR_j\} ; \quad k = \frac{2n\pi}{L} \quad (164)$$

It is customary to take both positive and negative values of k , so the values of $n \in [-N/2, N/2]$. This region in k -space is called the Brillouin Zone.

4.3.3 Transformation of the Energy

Using the Fourier transform discussed above, we can rewrite the microstate energy in terms of the normal modes.

$$E = \frac{1}{2}m \sum_k |\dot{x}_k|^2 + \frac{1}{2} \sum_k K(k) |x_k|^2 = \sum_k \left(\frac{1}{2}m |\dot{x}_k|^2 + \frac{1}{2}K(k) |x_k|^2 \right) \quad (165)$$

$$K(k) = K(1 - \exp\{ika\})(1 - \exp(-ika)) = 4K \sin^2(ka/2)$$

where $K = 2\pi/a$. The energy factorizes into N simple harmonic operators. $K(k)$ is the effective spring constant for the k -th SHO.

The frequency is given by the usual ratio

$$\omega(k) = \sqrt{\frac{K(k)}{m}} = 2\tilde{\omega}|\sin(ka/2)| \approx \omega_{\tilde{k}}a \quad (166)$$

where $\tilde{\omega} = \sqrt{K/m}$. The speed of sound is a constant $v(k) = a\tilde{\omega}$.

4.3.4 Classical vs Quantum and the Debye Approximation

The partition function and average energy for the classical model is given

$$Z_{class} = \prod_k (\beta\hbar\omega(k))^{-1} ; \quad U = 3Nk_B T \quad (167)$$

The second identity suggests a specific heat of $3k_B$, which is the Law of Dulong and Petit.

The quantum partition function and average energy is more complex

$$Z_{QM} = \prod_k \left[\frac{1}{2}\hbar\omega(k) + k_B T \ln(1 - \exp\{-\beta\hbar\omega(k)\}) \right]$$

$$U_{QM} \simeq \frac{L}{\pi} \int_{-\pi/a}^{\pi/a} dk \left[\frac{1}{2}\hbar\omega(k) + \frac{\hbar\omega(k)}{\exp\{\beta\hbar\omega(k)\} - 1} \right] \quad (168)$$

Typically, this integral is difficult to calculate analytically. This is especially difficult when we move to three dimensions and need to take lattice symmetry into account. Instead, we can apply the Debye approximation.

The Debye method uses the following approximations

1. We can use an approximate energy spectrum that is linear in the wave number, \mathbf{k} , and spherically symmetric.

$$\epsilon(\mathbf{k}) = \hbar v |\mathbf{k}| \quad (169)$$

2. The Brillouin Zone is spherically symmetric.
3. The Brillouin Zone contains exactly N k -values and $3N$ modes.

This allows us to simplify our complex 3D integral for the energy into a one-dimensional integral.

$$U_{Debye} = \frac{3\pi}{2} \int_0^n n^2 dn \frac{\hbar \omega(n)}{\exp\{\beta \hbar \omega(n)\} - 1} ; \quad n_D = \left(\frac{6N}{\pi} \right)^{1/3} \quad (170)$$

where n_D is the Debye radius. Furthering this calculation, we arrive at a simple integral as a function of Θ_D , the Debye temperature.

$$U_{Debye} = \frac{3L^3(k_B T)^4}{2\pi^2 \hbar^3 v^3} \int_0^{\Theta_D/T} dx x^3 (\exp\{x\} - 1)^{-1} \quad (171)$$

In the high and low T limits, the Debye energy is a good fit to the expected result.

$$U_{Debye}(T \gg \Theta_D) \approx 3Nk_B T ; \quad U_{Debye}(T \ll \Theta_D) \approx V \frac{\pi^2}{10\hbar^3 v^3} (k_B T)^4 \quad (172)$$

In the high temperature limit, U_{Debye} follows the Law of Dulong and Petit. In the low temperature limit, U_{Debye} looks like a black-body.

4.4 Ideal Quantum Gases

[Sw 308-325]

Consider a quantum ideal gas in a cubic box of volume $V = L^3$. The sides of the box are assumed to be impenetrable. The Hamiltonian of a single particle of mass m is just the kinetic energy term $H = -\frac{\hbar^2}{2m} \nabla^2$. The boundary conditions insist that the wave function vanishes at the edges of the box.

4.4.1 Quantum States

The single-particle wave function takes the form

$$\psi_{\mathbf{k}}(\mathbf{r}) = \sqrt{\frac{8}{L^3}} \sin(k_x x) \sin(k_y y) \sin(k_z z) ; \quad k_i = \frac{n_i \pi}{L} \quad (173)$$

where $n_i \in \mathbb{Z}$ and $i \in \{x, y, z\}$. The energy of a point in k -space is spherically symmetric. We can express it in terms of $\mathbf{n} = (n_x, n_y, n_z)$.

$$\epsilon_{\mathbf{k}} = \epsilon_{\mathbf{n}} = \frac{\hbar^2 \pi^2}{2mL^2} n^2 ; \quad n = |\mathbf{n}| \quad (174)$$

The density of states can be found by simply transforming variables to $\epsilon_{\mathbf{n}}$.

$$D(\epsilon) = \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \epsilon^{1/2} \quad (175)$$

An N -particle wave function can be constructed by multiplying single-particle wave functions. For distinguishable particles, this is straightforward

$$\psi_N = \prod_{j=1}^N \psi_{\alpha_j}(\mathbf{r}_j) \quad (176)$$

For identical particles, however, we must distinguish between bosons and fermions. For bosons, the wave function must be symmetrized. For fermions, the wave function must be anti-symmetrized.

The resultant wave functions are given by

$$\psi_N^{BE} = X_{BE} \sum_P P \left[\prod_{j=1}^N \psi_{\alpha_j}(\mathbf{r}_j) \right] ; \quad \psi_N^{FD} = X_{FD} \sum_P (-1)^P P \left[\prod_{j=1}^N \psi_{\alpha_j}(\mathbf{r}_j) \right] \quad (177)$$

where $P[\psi]$ are permutations of the exchange of the system's particles and X_{BE} and X_{FD} are normalization constants.

4.4.2 Grand Canonical Ensemble

We'll define the quantum grand canonical ensemble pdf analogously to the classical case.

$$P(E, N) = \frac{1}{\mathcal{Z}} \Omega(E, V, N) \exp\{-\beta E + \beta \mu N\} ; \quad \mathcal{Z} = \sum_{N=0}^{\infty} \sum_E \Omega(E, V, N) \exp\{-\beta E + \beta \mu N\} \quad (178)$$

For a system of non-interacting particles, the N -particle energies are given by $E = \sum_{j=1}^N \epsilon_{\alpha_j}$ where $\{\alpha_j\}$ is the set of quantum numbers for a given microstate. If each particle's Hamiltonian and quantum numbers are the same, we can describe the system by the occupation of each energy level instead of the quantum numbers of each particle $\{\alpha_j\} \rightarrow \{n_{\epsilon}\}$. Using this notation, the partition function can be written in a simple form.

$$\mathcal{Z} = \sum_{\{n_{\epsilon}\}} \prod_{\epsilon} \exp\{-\beta(\epsilon - \mu)n_{\epsilon}\} = \prod_{\epsilon} \sum_{n_{\epsilon}} \exp\{-\beta(\epsilon - \mu)n_{\epsilon}\} \quad (179)$$

Note, we can exchange the product and sum in the partition function expression. The sum $\sum_{\{n_{\epsilon}\}}$ is over all possible occupation numbers of the system, while the sum $\sum_{n_{\epsilon}}$ is over the possible occupation number for a given energy ϵ .

Due to the Pauli exclusion principle, all n_ϵ for fermions must be 0 or 1. For bosons, n_ϵ can take any value in the range $[0, \infty]$. Due to this distinction, we calculate different expected occupations.

$$\langle n_\epsilon \rangle = (\exp\{\beta(\epsilon - \mu)\} \pm 1)^{-1} \quad (180)$$

for fermions and bosons, respectively. These are the Fermi-Dirac and Bose-Einstein states, respectively.

The expectation values for N and U can thus be written as

$$\langle N \rangle = \sum_\epsilon \langle n_\epsilon \rangle ; \quad U = \langle E \rangle = \sum_\epsilon \epsilon \langle n_\epsilon \rangle \quad (181)$$

As a final note, we can write the pressure P of an (extensive) quantum gas. This works for both ideal Fermi and Bose gases.

$$P = \frac{2U}{3V} \quad (182)$$

4.5 Bose-Einstein and Fermi-Dirac Statistics

[Sw 326-350]

We will examine the impacts of the Bose-Einstein and Fermi-Dirac distributions.

4.5.1 Polylogarithms

First, a quick digression on polylogarithms. The polylogarithm $L_\nu(z)$ is defined in the following way

$$L_\nu(z) = \frac{1}{\Gamma(\nu)} \int_0^\infty dt \frac{t^{\nu-1}}{z^{-1}e^t - 1} ; \quad z \leq 1, \nu > 0 \quad (183)$$

where $\Gamma(\nu) = (\nu - 1)!$. Normally, this expression is not analytically solvable. We can also rewrite this formula in a more manageable form for $\nu > 1$.

$$L_\nu(z) = -\frac{1}{\Gamma(\nu - 1)} \int_0^\infty dt t^{\nu-2} \ln(1 - ze^{-t}) \quad (184)$$

A unique property of polylogarithms is their derivative follows the following identity.

$$z \frac{d}{dz} L_{\nu+1}(z) = L_\nu(z) \quad (185)$$

Below are some useful examples of polylogarithms

$$L_\nu(0) = 0 ; \quad L_\nu(1) = \zeta(\nu) ; \quad L_0(x) = \frac{x}{x-1} ; \quad L_1(x) = -\ln(1-x) \quad (186)$$

4.5.2 Bose-Einstein Statistics

Every occupation number must be positive $\langle n_\epsilon \rangle > 0$. Due to the Bose-Einstein distribution, we can then state that $\mu > \epsilon_0$. This takes a particularly nice case if we normalize the energies such that the ground state $\epsilon_0 = 0$. Then, the chemical potential of a gas of bosons must be negative.

$$\mu < 0 \quad (187)$$

We will now seek to find μ as a function of system parameters. Using the density of states $D(\epsilon)$ that we calculated previously, we can write

$$\begin{aligned} N &= \int_0^\infty D_{BE}(\epsilon) (\exp\{\beta(\epsilon - \mu)\} - 1)^{-1} d\epsilon \\ &= \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} (k_B T)^{3/2} \Gamma(3/2) L_{3/2}(\lambda) \end{aligned} \quad (188)$$

where $L_{3/2}(\lambda)$ is a polylogarithm and $\lambda = \exp\{\beta\mu\}$ is called the fugacity. From our discussion of polylogarithms, we know that this is only a valid expression for $\lambda \leq 1$ (outside of these bounds, the integrand diverges). This expression therefore cannot be valid below a temperature T_E that is determined by setting $L_{3/2}(\lambda)$ to its maximum value at $\lambda = 1$. We find an expression for T_E .

$$k_B T_E = \left(\frac{2\pi\hbar^2}{m} \right) \left(\frac{N}{2.612V} \right)^{2/3} \quad (189)$$

You can invert the above equation to find N . Below the Einstein temperature, the ground state occupation N_0 is significantly higher than the excited states. It is given by

$$N_0 = [\exp\{-\beta\mu\} - 1]^{-1} = N \left[1 - \left(\frac{T}{T_E} \right)^{3/2} \right] \quad (190)$$

Using the approximation that $\beta\mu$ is small below T_E , we arrive at an approximate value for μ .

$$\mu \approx -\frac{k_B T}{N} \left[1 - \left(\frac{T}{T_E} \right)^{3/2} \right]^{-1} \quad (191)$$

We can calculate values for U and c_V as well.

$$U = 1.7826 \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} (k_B T)^{5/2} ; \quad c_V = 1.925 k_B \left(\frac{T}{T_E} \right)^{3/2} \quad (192)$$

4.5.3 Fermi-Dirac Statistics

The Fermi Dirac distribution $\langle n_\epsilon \rangle$ takes an interesting form, shown in Figure 6. The distribution becomes a step function at $T = 0$. It crosses over from $1 \rightarrow 0$ at $\epsilon = \mu$. We will define the Fermi energy as

$$\epsilon_F = \lim_{T \rightarrow 0} \mu(T, N) \quad (193)$$

Note, the following useful identity

$$(\exp\{\beta(\epsilon - \mu)\} + 1)^{-1} = 1 - (\exp\{-\beta(\epsilon - \mu)\} + 1)^{-1} \quad (194)$$

For fermions of continuous (or almost continuous) energy spectra, we can integrate over the density of states $D_{FD}(\epsilon)$ to find N . Since the Fermi distribution becomes a step function as $T \rightarrow 0$, we can write N as the integral from $[0, \epsilon_F]$ using the density of states found in the previous section.

$$N = \int_0^{\epsilon_F} D_{FD}(\epsilon) d\epsilon = \frac{V}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \frac{2}{3} \epsilon_F^{3/2} \quad (195)$$

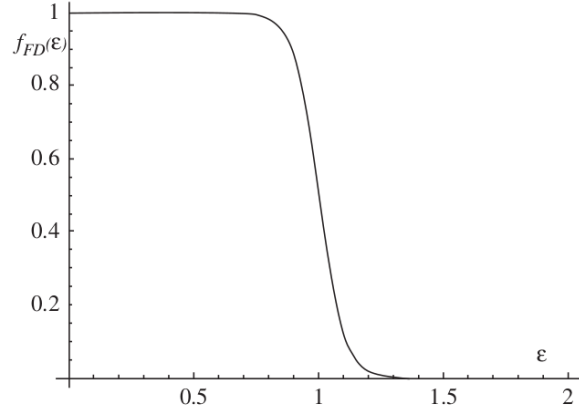


Figure 6: Form of $\langle n_\epsilon \rangle$ taken at $\mu = 1$ and $\beta = 10$.

Inverting this for ϵ_F , we arrive at an expression for the Fermi energy.

$$\epsilon_F = \frac{\hbar^2}{2m} \left(\frac{3\pi^2 N}{V} \right)^{2/3} ; \quad \epsilon_F = k_B T_F \quad (196)$$

where T_F is called the Fermi temperature.