

# Thermal Physics Notes

Charles Yang

Last updated: May 7, 2021



# Contents

<b>I</b>	<b>Entropy and Ideal Gases</b>	<b>1</b>
<b>1</b>	<b>Classical Ideal Gas</b>	<b>3</b>
<b>2</b>	<b>Discrete Random Variables</b>	<b>5</b>
2.1	Events and Random Experiments . . . . .	5
2.2	Discrete Random Variable . . . . .	6
2.3	Functions of Random Variables . . . . .	7
2.4	Binomial Distribution . . . . .	9
2.5	Multinomial Distribution . . . . .	11
2.6	Classical Gas . . . . .	12
<b>3</b>	<b>Continuous Random Variables</b>	<b>13</b>
3.1	Definitions . . . . .	13
3.2	Delta Function . . . . .	14
<b>4</b>	<b>Configurational Entropy</b>	<b>17</b>
4.1	Probability Distribution of Particles . . . . .	18
<b>5</b>	<b>Energy dependence of Entropy</b>	<b>21</b>
<b>6</b>	<b>Ideal Gas</b>	<b>25</b>
6.1	Indistinguishable Particles . . . . .	26
6.2	Entropy of interacting particles . . . . .	27
<b>7</b>	<b>Thermodynamic Variables</b>	<b>29</b>
7.1	Pressure . . . . .	29
7.2	Derivatives of Entropy . . . . .	31

7.3	Asymmetric Pistons . . . . .	32
7.4	Differential form of thermodynamics . . . . .	32
7.5	Measurements and Reservoirs . . . . .	32
<b>II</b>	<b>Postulates of Thermodynamics</b>	<b>33</b>
<b>8</b>	<b>Postulates and Laws of Thermodynamics</b>	<b>35</b>
8.1	Postulates . . . . .	35
8.2	Laws . . . . .	36
<b>9</b>	<b>Perturbations of State Functions</b>	<b>37</b>
9.1	Energy Conservation . . . . .	37
9.2	Integrating Factors in Classical Ideal Gas . . . . .	38
9.3	Work and Heat . . . . .	39
<b>10</b>	<b>Heat Engines and Refrigerators</b>	<b>41</b>
10.1	Heat Capacity . . . . .	41
10.2	Adiabatic Processes . . . . .	42
10.3	Thermodynamic Cycles . . . . .	43
10.4	Carnot Cycle . . . . .	44
<b>11</b>	<b>Thermodynamic Potentials</b>	<b>49</b>
11.1	Transforms of Energy . . . . .	50
<b>12</b>	<b>Consequences of Extensivity</b>	<b>53</b>
12.1	Euler Equation . . . . .	53
12.2	Gibbs-Duhem Relation . . . . .	54
12.3	Application to Ideal Gas . . . . .	54
<b>13</b>	<b>Thermodynamic Identities</b>	<b>57</b>
13.1	Second Derivatives . . . . .	57
13.2	Maxwell Relations . . . . .	57
13.3	“Insert $\partial(P, T)$ ”—Jacobians . . . . .	59
13.4	“Divide Through” . . . . .	60

13.5 Joule-Thomson Effect . . . . .	61
13.6 Heat Capacity . . . . .	62
13.7 General Strategy . . . . .	63
13.8 Examples . . . . .	63
<b>14 Extremum Principles</b>	<b>65</b>
14.1 Energy Minimum Principle . . . . .	65
14.2 Minimum Principle for Helmholtz Free Energy . . . . .	67
14.3 Minimum Principle for Enthalpy . . . . .	68
14.4 Minimum Principle for Gibbs Free Energy . . . . .	69
<b>15 Stability Conditions</b>	<b>71</b>
15.1 Extensive Stability Conditions . . . . .	71
15.2 Intensive Stability Conditions . . . . .	73
<b>16 Phase Transitions</b>	<b>75</b>
16.1 van der Waals Fluid . . . . .	75
16.2 Helmholtz Free energy . . . . .	81
16.3 Latent Heat . . . . .	81
16.4 Clausius-Clapeyron Equation . . . . .	82
16.5 Gibb's Phase Rule . . . . .	82
<b>17 Nernst Postulate</b>	<b>85</b>
17.1 Classical Ideal Gas Violation . . . . .	85
17.2 Consequences . . . . .	85
<b>III Classical Statistical Mechanics</b>	<b>87</b>
<b>18 Statistical Ensembles</b>	<b>89</b>
18.1 Microcanonical Ensemble . . . . .	89
18.2 Canonical Ensemble . . . . .	90
18.3 Partition Function and Free Energy . . . . .	94
18.4 Momentum-Independent Interactions . . . . .	96
18.5 Factorization of the Partition Function . . . . .	98

<b>19 Canonical Ensembles</b>	<b>103</b>
19.1 Grand Canonical Ensemble . . . . .	103
19.2 Classical Ideal Gas . . . . .	104
19.3 Other Classical Ensembles . . . . .	105
<b>20 Entropy Revisited</b>	<b>107</b>
20.1 From the Partition Functions . . . . .	107
20.2 Massieu Functions . . . . .	108
<b>21 Irreversibility in Thermodynamics</b>	<b>109</b>
 <b>IV Quantum Statistical Mechanics</b>	 <b>111</b>
<b>22 Quantum Mechanics</b>	<b>113</b>
<b>23 Quantum Canonical Ensemble</b>	<b>117</b>
23.1 Entropy . . . . .	118
23.2 Derivatives of Thermal Averages . . . . .	120
23.3 Merging fun! . . . . .	121
23.4 Factorization of the Partition Function . . . . .	121
<b>24 Black Body Radiation</b>	<b>125</b>
24.1 Resonator Cavity . . . . .	125
24.2 Quantization . . . . .	127
<b>25 Harmonic Solid</b>	<b>131</b>
25.1 Energy . . . . .	134
<b>26 Ideal Quantum Gas</b>	<b>137</b>
26.1 Fermions and Bosons . . . . .	137
26.2 Distribution Functions . . . . .	140
26.3 Density of States . . . . .	141
26.4 Model Systems . . . . .	142
<b>27 Bose-Einstein Condensation</b>	<b>147</b>

27.1 Macroscopic Occupation . . . . .	147
27.2 Condensation . . . . .	148
27.3 Properties of a BEC . . . . .	149
<b>28 Ideal Fermi Gases</b>	<b>151</b>
28.1 Free Particle . . . . .	152
28.2 Sommerfeld Expansion . . . . .	154
28.3 Thermodynamic Observables . . . . .	157
<b>29 Insulators and Semiconductors</b>	<b>161</b>
29.1 Bloch Theorem . . . . .	161
29.2 Schrödinger Equation . . . . .	163
29.3 Semiconductor Band Structure . . . . .	167
29.4 Doping . . . . .	169
<b>30 Interactions and the Ising Model</b>	<b>173</b>
30.1 Ising Model . . . . .	174
30.2 No Coupling . . . . .	175
30.3 Ising Chain . . . . .	176
30.4 2D Ising Square Lattice . . . . .	180
<b>31 Critical Phenomena</b>	<b>185</b>
31.1 Critical Exponents . . . . .	185
31.2 Landau Theory . . . . .	188
31.3 MFT Validity . . . . .	192
31.4 Critical Exponent Equalities . . . . .	193
31.5 Actual Values . . . . .	195





# Part I

## Entropy and Ideal Gases



# Chapter 1

## Classical Ideal Gas

The ideal gas will be used to introduce the idea of *entropy*. In an ideal gas, the only interaction between particles is elastic collisions; vdW effects and etc. are not considered. The ideal gas approximation is useful for gases at low densities.

In an ideal gas, the particles occupy a negligible volume compared to the volume of the gas; they are essentially point particles. These particles obey Newton's laws of motion, and collisions are elastic and take a negligible amount of time. The particles are allowed to collide with each other and the walls of the container.

For a collection of  $N$  particles in a volume  $V$ , each particle is characterized by a position  $\vec{r}_j$  and momentum  $\vec{p}_j$ , which each have 3 degrees of freedom. These positions and momenta form collections  $q = \{\vec{r}_j | j = 1, \dots, N\}$  and  $p = \{\vec{p}_j | j = 1, \dots, N\}$ . Together,  $q$  and  $p$  form the phase space, where points  $\{q, p\}$  indicate a specific microstate.

The kinetic energy of a particle is given  $E_j = \frac{\|\vec{p}_j\|^2}{2m}$ . Thus, the total energy of the system is given:

$$E = \sum_j^N \frac{\|\vec{p}_j\|^2}{2m} \quad (1.1)$$

There are no interactions between particles, so the potential energy  $U = 0$ . The particles are also considered to be indistinguishable.

Taking advantage of the fact that particles are considered to be indistinguishable, probability theory and statistics are used to examine the properties of the classical ideal gas. Naturally, assumptions on the distributions of positions and momenta will have to be made; these assumptions will be taken to be as simple as possible, and everything we don't know is considered to be equally likely. These assumptions will be validated by computing their consequences and comparing to experimental data.



# Chapter 2

## Discrete Random Variables

Boltzmann defined entropy in terms of the logarithm of the probability of a macroscopic state. The equation for entropy, written by Plank, is

$$S = k \log W \quad (2.1)$$

The key assumptions we will apply in using this equation is that the positions and momenta are independent (i.e. not correlated.). This results in the position and momentum contributions of entropy to be independent, and can be summed simply.

The concept of Entropy will be motivated using statistics.

### 2.1 Events and Random Experiments

A *sample space*  $\Omega$  is the set of all possible outcomes of a random experiment. An *event*  $A$  is a subset  $A \subseteq \Omega$ , which is a set of possible outcomes of the random experiment. An event with only one outcome is an *elementary event*. *Disjoint events*  $A, B \subseteq \Omega$  are events such that  $A \cap B = \{\}$ , or that are *mutually exclusive*.

A *probability*  $P$  is a function that gives the probability  $P(A)$  of event  $A$  occurring. This probability satisfies the following axioms:

1. For all  $A \subseteq \Omega$   $P(A) \geq 0$
2.  $P(\Omega) = 1$
3.  $P(A_1 \cup A_2 \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n)$

From these axioms, we can conclude  $0 \leq P(A) \leq 1$ ,  $P(\emptyset) = 0$ .

A *Conditional Probability*, of  $B$  given  $A$  is

$$P(B|A) \equiv \frac{P(A \cap B)}{P(A)} \quad (2.2a)$$

This is often rewritten as

$$P(A \cap B) \equiv P(B|A) \cdot P(A) \quad (2.2b)$$

Assume that  $A \subseteq B$  is a strict subset. Then, the following are true:

$$P(B) \geq P(A)$$

$$P(B - A) = P(B \setminus A) = P(B) - P(A)$$

Assume  $A, B$  are any two events. Then,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

This is related to the Principle of Inclusion/Exclusion (PIE)

There are two viewpoints of probability: *frequentist* and *Bayesian*. The former defines the probability as the asymptotic frequency of success of an event, while the latter defines probability as a description of a person's knowledge outcome of an experiment.

## 2.2 Discrete Random Variable

A *random variable* is the combination of a random event and its probability. If there is a finite or countable number of elementary events, the event is a *discrete* random variable.

Consider two events  $a$  and  $b$ . A *joint* probability  $P(a, b)$  can be defined. A joint probability refers to the probability that both event  $a$  and  $b$  occur. From the joint probability, the *marginal probability* can be defined:

$$P_A(a) = \sum_b P(a, b) \quad P_B(b) = \sum_a P(a, b)$$

The conditional probability is defined in a similar way to Equation 2.2b. The conditional and joint probabilities are related in the following manner:

$$P(a, b) = P(a|b) \cdot P_B(b) = P(b|a) \cdot P_A(a)$$

Bayes' theorem can also be stated:

$$P(a|b) = \frac{P(b|a)P_A(a)}{P_B(b)} \quad (2.3)$$

### 2.2.1 Independent and Dependent Variables

When two variables are considered *independent* iff  $P(a, b) = P_A(a)P_B(b)$ . For more than 2 variables, there is *pairwise* independence and *mutual* independence.

Dependent Variables: Marbles

You have a bag with 3 red marbles and 2 blue marbles. What are the probabilities when you take two marbles out of the bag?

Drawing the first marble changes the probability distribution of the second ball.

## 2.3 Functions of Random Variables

Given a random variable  $A = \{a_j | j = 1, \dots, N_A\}$ , a function  $F = \{F(a_j) | j = 1, \dots, N_A\}$  can be defined. The probability distribution of  $F$  is then

$$P_F(f) = \sum_a \delta_{f, F(a)} P_A(a)$$

The delta function is included because in general,  $F$  is not injective.

A common function of two variables is a sum. Let  $Z = X + Y$  be the sum of two random variables. The probability distribution of  $Z$  is then:

$$P_Z(z) = \sum_{x,y} \delta_{z, x+y} P(x, y)$$

Sum of Two Dice

You have two fair dice. The probability distribution of the sum of the two dice is

$$P_S(s) = \frac{1}{36} \sum_{x,y} \delta(s, x + y)$$

This distribution becomes:

Sum	2	3	4	5	6	7	8	9	10	11	12
Probability	1/36	1/18	1/12	1/9	5/36	1/6	5/36	1/9	1/12	1/18	1/36

### 2.3.1 Properties of Functions

The *expected value* or *mean* of a function is given:

$$\langle F \rangle = \bar{F} \equiv \sum_a F(a) P_A(a) \quad (2.4)$$

The  $n$ -th moment of  $F$  is

$$\langle F^n \rangle \equiv \sum_a F(a)^n P_A(a) \quad (2.5)$$

The variance is given:

$$\sigma_F^2 = \langle (F - \langle F \rangle)^2 \rangle = \langle F^2 \rangle - \langle F \rangle^2 \quad (2.6)$$

and the standard deviation is the square root of the variance:

$$\sigma_F \equiv \sqrt{\langle F^2 \rangle - \langle F \rangle^2} \quad (2.7)$$

A correlation function  $f_{FG}$  is defined:

$$f_{FG} = \langle FG \rangle - \langle F \rangle \langle G \rangle \quad (2.8)$$

If  $F$  and  $G$  are independent, then the correlation function is naturally zero, i.e. they are not correlated.

### 2.3.2 Properties of Expected Value

The expected value of a sum is:

$$\langle X + Y \rangle = \langle X \rangle + \langle Y \rangle \quad (2.9)$$

From this, we can show that the variance of a sum is given:

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2 \text{Cov}(X, Y) \quad (2.10)$$

where the covariance is defined as

$$\text{Cov}(X, Y) = \langle XY \rangle - \langle X \rangle \langle Y \rangle \quad (2.11)$$

### 2.3.3 Sets of Independent Variables

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

$$S = \sum_j F_j$$

The expected value of the sum follows from Equation 2.9

$$\langle S \rangle = \sum_j \langle F_j \rangle$$

Further, the variance of  $S$  is:

$$\begin{aligned} \sigma_S^2 &= \langle S^2 \rangle - \langle S \rangle^2 = \sum_{jk} \langle F_j F_k \rangle - \langle F_j \rangle \langle F_k \rangle \\ &= \sum_{j \neq k} \langle F_j F_k \rangle + \sum_j \langle F_j^2 \rangle - \sum_{j \neq k} \langle F_j \rangle \langle F_k \rangle - \sum_j \langle F_j \rangle \langle F_j \rangle \end{aligned}$$



$$\begin{aligned}
&= \sum_{j \neq k} \langle F_j \rangle \langle F_k \rangle + \sum_j \langle F_j^2 \rangle - \sum_{j \neq k} \langle F_j \rangle \langle F_k \rangle - \sum_j \langle F_j \rangle^2 \\
&= \sum_j \langle F_j^2 \rangle - \langle F_f \rangle^2 \\
&= \sum_j \sigma_j^2
\end{aligned}$$

For the case that  $F_j$  have the same mean and variance, the standard deviation of  $S$  grows with  $\sigma_S = \sigma_F \sqrt{N}$ . However, the relative standard deviation increases:

$$\frac{\sigma_S}{\langle S \rangle} = \frac{\sigma_F \sqrt{N}}{N \langle F \rangle} = \frac{\sigma_F}{\langle F \rangle \sqrt{N}} \quad (2.12)$$

## 2.4 Binomial Distribution

### 2.4.1 Combinatorics

If we have  $N$  objects, the number of ways we can order the objects is  $N!$ . This is because there are  $N$  choices for the first object,  $N - 1$  choices for the second, etc. If we truncate the ordering after  $k$  objects, there are instead  $N!/(N - k)!$  orderings. If we do not care about the order of the objects—i.e., we want to choose  $k$  out of  $N$  objects—the number of ways to do so is:

$$\binom{N}{k} = \frac{N!}{(N - k)! k!} \quad (2.13)$$

### Binomial Distribution

We have  $N$  variables which take the value 1 with probability  $p$  and 0 with probability  $(1 - p)$ . It is easy to see that:

$$\begin{aligned}
\langle F \rangle &= 1p + 0(1 - p) = p \\
\langle F^2 \rangle &= p \\
\sigma^2 &= p(1 - p)
\end{aligned}$$

The distribution of the sum is:

$$\begin{aligned}
\langle S \rangle &= pN \\
\sigma_S^2 &= p(1 - p)N \\
\frac{\sigma_S}{\langle S \rangle} &= \sqrt{\frac{1 - p}{pN}}
\end{aligned}$$

The probability that  $0 \leq n \leq N$  of the variables is 1 is given

$$p^n (1 - p)^{N - n}$$

However, there are  $\binom{N}{n}$  that subsets of size  $n$  can be chosen. Thus, the binomial distribution, denoted  $P_B(n|N)$  is given:

$$P_B(n|N) = \binom{N}{n} p^n (1-p)^{N-n} \quad (2.14)$$

The binomial distribution get its name from the binomial theorem of how to expand binomials:

$$[p + (1-p)]^N = \sum_{n=0}^N \binom{N}{n} p^n (1-p)^{N-n}$$

For a very large  $N \rightarrow \infty$ , the binomial distribution can be represented by a Gaussian.

$$g(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[ -\frac{(x-x_0)^2}{2\sigma^2} \right] \quad (2.15)$$

The mean and variance are as expected, with

$$\langle x \rangle = x_0 \quad \sigma_x^2 = \sigma^2$$

The binomial distribution can then be written:

$$p_B(n|N) \approx \frac{1}{\sqrt{2\pi p(1-p)N}} \exp \left[ -\frac{(n-pN)^2}{2p(1-p)N} \right] \quad (2.16)$$

### 2.4.2 Sterling's Approximation

In the simplest version, Sterling's approximation gives the logarithm of the factorial:

$$\begin{aligned} \ln N! &= \sum_{n=1}^N \ln(n) \\ &\approx \int_1^N dx \ln x \\ &= N \ln N - N + 1 \end{aligned} \quad (2.17)$$

More elaborately comes from the gamma function. At large  $N$ , the integrand is a sharp peak with a maximum at  $x = N$  and value  $N^N e^{-N}$ .

$$\begin{aligned} N! &= \int_0^\infty e^{-x} x^N dx \\ &\approx \int_0^\infty N^N e^{-N} \exp \left[ -\frac{(x-N)^2}{2\sigma^2} \right] dx \end{aligned} \quad (2.18)$$

The value of  $\sigma$  is determined by matching the second derivative of  $\ln(g)$  at the peak, which gives  $\sigma^2 = N$

$$= \int_0^\infty N^N e^{-N} \exp \left[ -\frac{(x-N)^2}{2N} \right]$$

$$\begin{aligned}
&\approx \int_{-\infty}^{\infty} dx N^N e^{-N} \\
&= N^N e^{-N} \sqrt{2\pi N}
\end{aligned} \tag{2.19}$$

Finally, is Gosper's approximation:

$$N! \approx N^N e^{-N} \sqrt{\left(2N + \frac{1}{3}\right) \pi} \tag{2.20}$$

Only the first Sterling approximation is necessary, as the error is of the order  $\frac{1}{N}$ , which, for quantities on the order of a mole, is  $\sim 10^{-23}$ .

Using the first Sterling approximation, the log of the binomial distribution (at large  $N$ ) becomes:

$$\begin{aligned}
\ln P_B(n|N) &\approx N \ln N - N - n \ln n + n - (N - n) \ln(N - n) \\
&\quad + (N - n) + n \ln p + (N - n) \ln(1 - p) \\
&= N \ln N - n \ln n - (N - n) \ln(N - n) + n \ln p + (N - n) \ln(1 - p)
\end{aligned} \tag{2.21}$$

Thus, we can consider the binomial distribution as a continuous variable in  $n$  to find the peak. After a little bit of calculus, the maximum is at  $n_0 = pN$ , which agrees with the exact result. Further, the expected value of the distribution is at the maximum.

## 2.5 Multinomial Distribution

The binomial distribution can be viewed as putting  $N$  objects into 2 boxes, with probability  $p$  and  $1 - p$  respectively. If we instead generalize the problem to  $M$  boxes, with probabilities  $p_1, \dots, p_M$ , we obtain the *multinomial distribution*.

The probability of a particular subset  $N_j$  particles being assigned to box  $j$  is

$$\prod_{j=1}^M p_j^{N_j} \tag{2.22}$$

and the probabilities are normalized with

$$\sum_{j=1}^M p_j = 1 \tag{2.23}$$

and sets partitioned with

$$\sum_j N_j = N \tag{2.24}$$

If the objects are distinct, the ways that the balls can be divided into  $M$  boxes is:

$$\frac{N!}{\prod_{j=1}^M N_j} \quad (2.25)$$

Note that this is equivalent to permutations of a string of characters with repeats, such as “aaabbcccd,” as each letter would represent a box, and the position of the letter represents the distinct particle.

## 2.6 Classical Gas

For a classical gas, we can assume the probability of finding a particular particle in a volume  $V_j$  to be proportional to  $V_j$  (think geometric probability). Thus, the probability becomes:

$$p_j = \frac{V_j}{V} \quad (2.26)$$

That is, the probability density is uniform throughout the total volume (no areas are more probable than others). This means, that the distribution of a classical ideal gas is:

$$P_M(N_j) = \frac{N!}{V^N} \prod_{j=1}^M \left( \frac{V_j^{N_j}}{N_j!} \right) \quad (2.27)$$

# Chapter 3

## Continuous Random Variables

### 3.1 Definitions

Previously we considered discrete variables, such as a die, which can only take integer values from 1 – 6. If we instead consider a *continuous* random process that can produce any real number between 0 – 6, we can no longer consider the value of a single outcome, as it becomes  $P = 1/\infty = 0$ . Instead, we take an integral:

$$P([a, b]) = A \int_a^b dx = (b - a)A$$

where  $A$  is a normalization constant. For the continuous die, the normalization constant is  $A = \frac{1}{6}$ . Similar to how discrete variables can have different probabilities, a *probability density* can be ascribed to each point. Thus, instead, we have for a probability density:

$$P([a, b]) = \int_a^b P(x) dx \tag{3.1}$$

Dimensionally, probability densities are unitless, but probability densities *do* have densities—for example, per volume, or per length. Further, probabilities must be normalized over the entire sample space:

$$1 = \int_{\Omega} P(x) dx \tag{3.2}$$

Marginal probabilities are then defined:

$$P_x(x) = \int_{-\infty}^{\infty} P(x, y) dy \tag{3.3}$$

Conditional Probabilities:

$$P(y|x) = \frac{P(x, y)}{P_x(x)} \tag{3.4a}$$

$$P(x, y) = P(x|y)P_y(y) = P(y|x)P_x(x) \tag{3.4b}$$

Which can be rearranged to form Bayes's Theorem.

Indpendence maintains its definition as well:

$$P(x, y) = P_x(x)P_y(y) \quad (3.5)$$

The expected value of a function, rather than being a discrete sum, becomes an integral:

$$\langle F \rangle = \int_{\Omega} F(x)P(x) \, dx \quad (3.6)$$

And similarly with the moments and central moments:

$$\langle x^n \rangle = \int_{\Omega} x^n P(x) \, dx \quad (3.7a)$$

$$\langle (x - \langle x \rangle)^n \rangle = \int_{\Omega} (x - \langle x \rangle)^n P(x) \, dx \quad (3.7b)$$

## 3.2 Delta Function

Finally, we define the Dirac delta to determine the sum of two variables. The delta function is defined with the following properties:

$$\delta = \lim_{\varepsilon \rightarrow 0} \delta_{\varepsilon}, \quad \delta_{\varepsilon}(x) \equiv \begin{cases} 0 & x < -\varepsilon \\ \frac{1}{2\varepsilon} & -\varepsilon \leq x \leq \varepsilon \\ 0 & \varepsilon < x \end{cases} \quad (3.8a)$$

Normalization

$$\int_{-\infty}^{\infty} \delta(x) \, dx = 1 \quad (3.8b)$$

$$\int_a^b \delta(x) \, dx = 1 \iff x \in (a, b) \quad (3.8c)$$

Other properties are:

$$\delta(x) = \delta(-x)$$

$$\delta(cx) = \delta(|c| x)$$

$$\int_a^b \delta(cx) \, dx = \int_{a/|c|}^{b/|c|} \delta(y) \frac{dy}{|c|} = \frac{1}{|c|}$$

### 3.2.1 Integrating functions with $\delta(x)$

The definition of the delta function indicates that:

$$\int_a^b f(x)\delta(x - x_0) = \begin{cases} f(x_0) & x_0 \in (a, b) \\ 0 & \text{otherwise} \end{cases}$$

It can be shown that this is true using a Taylor Series expansion of  $f$  about  $x_0$  with  $\delta_\epsilon$  as  $\epsilon \rightarrow 0$ .

This can be extended to

$$\int_{-\infty}^{\infty} f(x)\delta(c(x - x_0)) \, dx = \frac{f(x_0)}{|c|} \quad (3.9)$$

Most generally, we want to be able to solve

$$\int_{-\infty}^{\infty} f(x)\delta \circ g(x) \, dx$$

The following applies when  $g$  has simple zeros (no algebraic multiplicity) and has a derivative  $g'(x)$ .

$$\int_a^b f(x)\delta \circ g(x) \, dx = \sum_{j=1}^n \frac{f(x_j)}{|g'(x_j)|} \quad (3.10)$$

where  $\{x_j | x_j \in (a, b) \text{ and } g(x_j) = 0\}$ .

Given two continuous variables  $x, y$ , we can define a new variable  $s = f(x, y)$ . Then, the probability distribution of  $s$  is:

$$P(s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(x, y)\delta(s - f(x, y)) \, dx \, dy \quad (3.11)$$

Note that this probability density is automatically 1.

Continuous Dice

Imagine two variable  $x, y$  which take values  $[0, 6]$  continuously with equal probability

$$P(x, y) = \frac{1}{36}$$

Then,

$$\begin{aligned} P(s) &= \int_0^6 \int_0^6 P(x, y) \delta(s - x - y) \, dx \, dy \\ &= \frac{1}{36} \int_0^6 \int_0^6 \delta(s - x - y) \, dx \, dy \end{aligned}$$

integrating wrt  $y$  first,

$$\int_0^6 \delta(s - x - y) \, dy = \begin{cases} 1 & 0 \leq s - x \leq 6 \\ 0 & \text{otherwise} \end{cases}$$

Then, integrating over  $x$ , we get an additional constraint



# Chapter 4

## Configurational Entropy

Boltzmann defined entropy of a composite system to be the natural logarithm of the probability of extensive variables, up to an additive and multiplicative constant. The essential property of entropy is that when two systems come into contact, the final entropy is the maximal entropy of the system.

The initial state of a system is usually an improbable state; when mixing coffee and milk, the initial state of separate coffee and milk phases is not as probable as a mixed state.

### 4.0.1 Assumptions

We assume that the position and momenta are independent of each other. This statement allows us to separate:

$$P(q, p) = P(q)P(p) \quad (4.1)$$

The part of the entropy arising from the positions  $q$  is the *configurational entropy*, and is dependent only on the volume and number of particles:

$$S_q = S_q(V, N)$$

The contribution of entropy from the momenta  $p$  is dependent only on the energy and the number of particles:

$$S_p = S_p(E, N)$$

The total entropy ends up being:

$$S_{tot}(E, V, N) = S_1(V, N) + S_p(E, N)$$

#### Intensivity and Extensivity

An extensive property depends on the amount of stuff; the volume, number of particles, and total energy are extensive properties. An intensive property does not depend on the amount of stuff, such as temperature, or density.

## 4.1 Probability Distribution of Particles

Imagine a composite system of two boxes; box  $J$  and box  $K$ . The total number of particles in the system is:

$$N_T = N_j + N_k$$

and the total volume is

$$V_T = V_j + V_k$$

Remove the partition between the boxes so the particles can exchange between the boxes. The number of particles in each box then becomes a function of time, that is:

$$N_T = N_j(t) + N_k(t)$$

The functions  $N_j$  and  $N_k$  will evolve until the macroscopic state with maximal probability is obtained.

From this, we make the simplest assumption: the particle positions are mutually independent:

$$P(q) = \prod_i P_i(r_i)$$

Further, we assume that all infinitesimal volumes are equally likely; the probability that a particle is in a given volume is given by the geometric probability of that volume relative to the entire volume.

We can then reduce the probability distribution of  $N_j, N_k$  to the binomial distribution

$$P(N_j, N_k) = \underbrace{\frac{N_T!}{N_j!N_k!}}_{\binom{N_T}{N_j}} \cdot \underbrace{\left(\frac{V_j}{V_T}\right)^{N_j}}_{p_j^{N_j}} \cdot \underbrace{\left(\frac{V_k}{V_T}\right)^{N_k}}_{(1-p_j)^{N_T-N_j}} \quad (4.2)$$

Thus, the mean of  $N_j$  is given:

$$\langle N_j \rangle = \frac{N_T V_j}{V_T} \quad (4.3a)$$

$$\sigma_{N_j}^2 = \frac{N_T V_j}{V_T} \left(1 - \frac{V_j}{V_T}\right) = \frac{N_T V_j V_k}{V_T^2} \quad (4.3b)$$

Similar expressions can be found for  $N_k$ . From this, we find that

$$\frac{\langle N_j \rangle}{V_j} = \frac{N_T}{V_T} = \frac{\langle N_k \rangle}{V_k} \quad (4.4)$$

The relative standard deviation

$$\frac{\sigma_{N_j}}{\langle N_j \rangle} = \sqrt{\frac{1}{\langle N_j \rangle}} \cdot \sqrt{\frac{V_k}{V_T}}$$

is small for large values of  $N$ ; for mole-like quantities, this will be on the order of  $10^{-10}$ .

Actual vs Expected Value It becomes tiresome to constantly write angle brackets around quantities, so macroscopic variables will typically represent expected values rather than their actual values. This is particularly important to remember when computing functions of these variables, such as entropy.

However, when discussing probabilities, we instead are using the actual value of the variable.

We will introduce a new function  $\Omega$  defined to be:

$$\Omega_q(N, V) = \frac{V^N}{N!} \quad (4.5)$$

This then allows us to rewrite the probability 4.2 as:

$$P(N_j, N_k) = \frac{\Omega_q(N_j, V_j)\Omega_q(N_k, V_k)}{\Omega(N_T, V_T)} \quad (4.6)$$

Because the logarithm is monotonic, the maximum of the probability is the maximum of the logarithm. Using the logarithm, we have:

$$\begin{aligned} \ln(P) &= \ln[\Omega_q(N_j, V_j)] + \ln[\Omega_q(N_k, V_k)] - \ln[\Omega_q(N_T, V_T)] \\ k \ln(P) &= k \ln[\Omega_q(N_j, V_j)] + cN_j + k \ln[\Omega_q(N_k, V_k)] + cN_k - k \ln[\Omega_q(N_T, V_T)] - cN_T \end{aligned}$$

The expression for  $k \ln(P)$  is separated separated into terms for each individual box, as well as the total box. In fact, the function  $k \ln[\Omega_q(N, V)] + cN$  looks like a good candidate for the entropy function, which has both the multiplicative and additive constants that Boltzmann described.

Thus, we assume that we can write the configurational entropy as:

$$S_q(N, V) = k \ln \left( \frac{V^N}{N!} \right) + kXN \quad (4.7)$$

for some constants  $k$  and  $X$ .

Then, the sum of the entropies of the two subsystems is

$$S_q(N_j, V_j) + S_q(N_k, V_k) = k \ln[P(N_j, N_k)] + S_q(N_T, V_T)$$

Because  $S_q(N_T, V_T)$  is constant, we need only maximize the first term. Using a gaussian approximation, we note that the first term is negligible when compared to the total configurational entropy; thus we can show that when maximised, the following relation holds:

$$S_q(N_j, V_j) + S_q(N_k, V_k) = +S_q(N_T, V_T) \quad (4.8)$$

This demonstrates the crucial property of *additivity* of entropy.

Generalizing to a system of  $M \geq 2$  subsystems, the probability becomes:

$$P(\{N_j|j = 1, \dots, M\}) = \frac{\prod_j^M \Omega(N_j, V_j)}{\Omega(N_T, V_T)}$$

$$k \ln[P(\{N_j|j = 1, \dots, M\})] = -k \ln \Omega_q(N_T, V_T) + \sum_j^M k \ln \Omega(N_j, V_j) \quad (4.9)$$

Once again, after maximising and discarding negligible terms, we obtain:

$$S_q(N_T, V_T) = \sum_j^M S_q(N_j, V_j) \quad (4.10)$$

Using the sterling approximation  $\ln(N!) \approx N \ln N - N$ , we obtain the configurational entropy:

$$S_q(N, V) = kN \left[ \ln \left( \frac{V}{N} \right) + X \right] \quad (4.11)$$

# Chapter 5

## Energy dependence of Entropy

We use the same system as in the previous chapter, but rather than allowing particles to transfer, we only allow energy to transfer; there is a *diathermal barrier* between the boxes. We assume that the only interactions between particles are elastic collisions. Further, we assume that every particle has the same mass  $m$ .

The energy of box  $\alpha$  is then:

$$E_\alpha = \sum_{i=1}^{N_\alpha} \frac{p_{\alpha,i}^2}{2m} \quad (5.1)$$

and is constrained by

$$E_T = \sum_{\alpha} E_\alpha$$

Further, all volumes and particle numbers stay constant.

We assume the momentum has a constant distribution, subject to constraints by the total energy of the system; all momenta are equally likely. The probability distribution of the energies is:

$$P(E_j, E_k) = \frac{\int_{-\infty}^{\infty} \delta \left( E_j - \sum_i^{N_j} \frac{p_{j,i}^2}{2m} \right) dp_j \cdot \int_{-\infty}^{\infty} \delta \left( E_k - \sum_i^{N_k} \frac{p_{k,i}^2}{2m} \right)}{\int_{-\infty}^{\infty} \delta \left( E_T - \sum_i^{N_j} \frac{p_{j,i}^2}{2m} - \sum_{i'}^{N_k} \frac{p_{k,i'}^2}{2m} \right) dp_j dp_k} \quad (5.2)$$

where the shorthand for integrals is given

$$\int_{-\infty}^{\infty} dp_\alpha \equiv \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} d^3 p_{\alpha,1} \dots d^3 p_{\alpha,N_\alpha}$$

To simplify notation, we define  $\Omega_p(E_\alpha, N_\alpha)$

$$\Omega_p(E_\alpha, N_\alpha) = \int_{-\infty}^{\infty} \delta \left( E_\alpha - \sum_i^{N_\alpha} \frac{p_{\alpha,i}^2}{2m} \right) dp_\alpha \quad (5.3)$$

reducing the probability to

$$P(E_j, E_k) = \frac{\Omega_p(E_j, N_j)\Omega_p(E_k, N_j)}{\Omega_p(E_T, N_t)} \quad (5.4)$$

Notice that we rewrite the sum of momenta in Equation 5.3 as a the norm of some  $3N$  dimensional vector. Then, the argument of the delta function becomes:

$$\left\| \vec{P} \right\| = \sqrt{2mE} \quad (5.5)$$

Thus, the probability becomes a geometric probability on the surface of a hypersphere. While we don't know the the explicit formula for the area of a hypersphere, we can determine its form using dimensional analysis:

$$A = S_n r^{n-1}$$

allowing us to rewrite:

$$\Omega_p(E, N) = \int_0^\infty S_{3N} p^{3N-1} \delta\left(E - \frac{p^2}{2m}\right) dp \quad (5.6)$$

using the change of variable

$$S_{3N} p^{3N-1} dp = \prod_i^N d^3 p_i$$

Using another change of variable  $p^2 = 2mu$ ,

$$\begin{aligned} \Omega_p(E, N) &= \int_0^\infty S_{3N} p^{3N-2} \delta\left(E - \frac{p^2}{2m}\right) p dp \\ &= \int_0^\infty S_{3N} (2mu)^{3N-2} \delta(E - u) m du \\ &= S_{3N} m (2Me)^{(3N-2)/2} \end{aligned}$$

Using gaussian integrals, we can determine  $S_n$  as:

$$\begin{aligned} \left[ \int_{-\infty}^\infty e^{-x^2} dx \right]^n &= \int_0^\infty \exp -r^2 S_n r^{n-1} dr \\ \pi^{n/2} &= \frac{1}{2} S_n \int_0^\infty e^{-t} t^{n/2-1} dt \\ &= \frac{1}{2} S_n \Gamma(n/2) \\ \implies S_n &= \frac{2\pi^{n/2}}{\Gamma(n/2)} \end{aligned} \quad (5.7)$$

Plugging this into our expression for  $\Omega_p$ :

$$\Omega(E, N) = \frac{(2\pi m)^{3N/2}}{\Gamma(3N/2)} E^{3N/2-1} \quad (5.8)$$

Taking the logarithm,

$$\ln \Omega_p(E, N) = \frac{3N}{2} \ln(2\pi n) + \left( \frac{3N}{2} - 1 \right) \ln E - \ln \Gamma(3N/2) \quad (5.9)$$

For very large  $N$ , we can approximate  $\Gamma(N) \sim N!$ , so

$$\ln \Gamma(3N/2) \approx \ln \left[ \left( \frac{3N}{2} - 1 \right)! \right] \approx \frac{3N}{2} \ln \frac{3N}{2} - \frac{3N}{2}$$

Thus,

$$\Omega_p(E, N) \approx N \left[ \frac{3}{2} \ln \frac{E}{N} + X \right]$$

Finally, we can rewrite the probability to find that it is proportional to:

$$P(E_j, E_T - E_k) \propto E_j^{3N_j/2-1} (E_T - E_j)^{3(N_T-N_j)/2-1} \quad (5.10)$$

We approximate the expected value of  $\langle E_j \rangle$  as the peak of the probability:

$$\begin{aligned} 0 &= \frac{\partial}{\partial E_j} \ln P(E_j, E_T - E_j) \\ &= \left( \frac{3N_j}{2} - 1 \right) \frac{1}{E_j} - \left( \frac{3(N_T - N_j)}{2} - 1 \right) \frac{1}{E_T - E_j} \\ \Rightarrow \langle E_j \rangle &\approx E_{j,\max} = \left( \frac{3N_j - 2}{3N_T - 4} \right) E_T \approx \frac{N_j}{N_T} E_T \end{aligned} \quad (5.11)$$

Or, more illuminatingly,

$$\frac{\langle E_j \rangle}{N_j} = \frac{E_T}{N_T} \quad (5.12)$$

To show that our approximations were valid, we look at the standard deviation. We approximate the distribution as a gaussian, and find that

$$-\frac{1}{\sigma_{E_j}^2} = \frac{\partial^2}{\partial E_j^2} \ln(P) \Big|_{E_j=E_{j,\max}}$$

Computing, this results in:

$$\sigma_{E_j} = \langle E_j \rangle \sqrt{\frac{1}{N_j}} \sqrt{\frac{2(N_T - N_j)}{3N_T}} \quad (5.13)$$

The relative width is

$$\frac{\sigma_{E_j}}{\langle E_j \rangle} = \sqrt{\frac{1}{N_j}} \sqrt{\frac{2(N_T - N_j)}{3N_T}}$$

Finally, we get that:

$$S_p(E, N) = k \ln \Omega_p(E, N) + cN$$

Using Sterling's approximation, we find that

$$S_p(E, N) = kN \left[ \frac{3}{2} \ln \left( \frac{E}{N} \right) + X \right] \quad (5.14)$$

It can be verified similar to configurational entropy that this expression satisfies the desired properties of entropy.



# Chapter 6

## Ideal Gas

Combining the results from the previous two chapters, the entropy of the classical ideal gas is given:

$$S(E, V, N) = kN \left[ \frac{3}{2} \ln \left( \frac{E}{N} \right) + \ln \left( \frac{V}{N} \right) + X \right] \quad (6.1)$$

From quantum mechanics, we can determine the value of  $X$  as:

$$X = \frac{3}{2} \ln \left( \frac{4\pi m}{3h^2} \right) + \frac{5}{2} \quad (6.2)$$

We know that entropy is maximized when a system reaches equilibrium. To verify our equation satisfies this property, we examine what happens when two systems at energy  $E_j$  and  $E_k$  are placed into contact. The entropy of the total system is given:

$$S_T = S(E_j, V_j, N_j) + S(E_k, V_k, N_k)$$

Maximising the entropy,

$$\begin{aligned} 0 &= \frac{\partial}{\partial E_j} (S(E_j, V_j, N_j) + S(E_T - E_j, V_j, N_j)) \\ &= \frac{\partial S_j}{\partial E_j} + \frac{\partial E_j}{E_k} \frac{\partial S_k}{\partial E_j} \\ &= \frac{\partial S_j}{\partial E_j} - \frac{\partial S_k}{\partial E_k} \end{aligned}$$

This yields the relation:

$$\frac{\partial S_j}{\partial E_j} = \frac{\partial S_k}{\partial E_k} \quad (6.3)$$

explicitly evaluating these derivatives, we get:

$$\frac{\partial S_j}{\partial E_j} = \frac{3kE_j}{2N_j} \implies \frac{E_j}{N_j} = \frac{E_k}{N_k} \quad (6.4)$$

which matches Equation 5.12

Similarly, if we put two open systems in contact, we get the condition

$$\frac{\partial S_j}{\partial N_j} = \frac{\partial S_k}{\partial N_k} \quad (6.5)$$

Evaluating,

$$\frac{\partial S_j}{N_j} = \frac{3}{2} \ln \frac{E_j}{N_j} + \ln \frac{V_j}{N_j} \quad (6.6)$$

Because the energy is also being exchanged, the condition 6.4 must be maintained. Thus, we obtain:

$$\frac{V_j}{N_j} = \frac{V_k}{N_k} \quad (6.7)$$

Finally, we test a system with a movable diathermal barrier. Taking the derivative wrt  $V_j$ ,

$$\frac{\partial S_j}{\partial V_j} = \frac{\partial S_k}{\partial V_j} \quad (6.8)$$

Evaluating, the equilibrium condition is found to be

$$\frac{N_j}{V_j} = \frac{N_k}{V_k}$$

which is the same condition as Equation 6.7!

Assuming that all configurations are equally likely (normalized by a factor  $Y$ ), the probability distribution of  $V_j$  is:

$$P(V_j) = Y V_j^{N_j} V_k^{N_k} \quad (6.9)$$

Maximising this probability distribution, we once again obtain the condition 6.7, which supports that entropy is maximised in the most probable macrostate.

To validate our approximations, we can test our predictions against test cases. While it is impossible to create an ideal gas, we can check using a dilute gas, which approximates a classical ideal gas. Alternatively, we can write a molecular dynamics simulation and evolve it according to newtonian mechanics.

## 6.1 Indistinguishable Particles

We have previously assumed that every particle in the distinguishable; however in quantum mechanics, particles of the same type are considered *indistinguishable*. Revisiting the entropy calculation under the assumption the particles are indistinguishable actually yields the same result!

For distinguishable particles, a permutation yields a different microscopic state. However, for indistinguishable particles, the permutation yields an identical microstate.

In the indistinguishable case, we determined the probability distribution to be

$$P(N_j, N_k) = \binom{N_T}{N_j} \left( \frac{V_j}{V_t} \right)^{N_j} \left( 1 - \frac{V_j}{V_T} \right)^{N_T - N_j}$$

We must replace the combinatoric with a 1, as each microstate is indistinguishable. Further, we must adjust the probabilities. Assuming that each position is equally likely, we neglect the  $y, z$  component and compute the  $x$  component. WLOG, we order the particles from largest to lowest  $x$  component. The probability distribution becomes:

$$P(N_j, N_k) = Y A_j^{N_j} A_k^{N_k} \int_0^{L_j} dx_1^j \int_0^{x_1^j} dx_2^j \dots \int_0^{x_{N_j-1}^j} dx_{N_j}^j \\ \times \int_0^{L_k} dx_1^k \int_0^{x_1^k} dx_2^k \dots \int$$

## 6.2 Entropy of interacting particles

Real gases have interactions between particles. The Hamiltonian of the system becomes:

$$H_N = \sum_i^N \frac{p_i^2}{2m} + \sum_i^N \sum_{i' > i}^N \phi(r_i, r_{i'}) \quad (6.10)$$

We can rewrite the Hamiltonian as the sum of the Hamiltonians of the two boxes, with an interaction Hamiltonian:

$$H_{N_T} = H_{N_j} + H_{N_k} + H_{jk} \quad (6.11)$$

where

$$H_{jk} = \sum_{i=1}^{N_j} \sum_{l=1}^{N_k} \phi(r_i, r_l) \quad (6.12)$$

Because the interaction distance is very short tane, the contribution to the Hamiltonian  $H_{jk}$  is significant only for particles near the interface between the boxes. The energy associated scales with  $V_N^{2/3} \sim A$ . Similarly, this requiers  $N_\alpha^{2/3}/N_\alpha$  of the particles to interact. Thus, the hamiltonian scalse with  $H_{jk} \propto N_j^{-1/3} N_k^{-1/3}$ , which can be neglected<sup>1</sup>.

Thus, the probability distribution is given:

$$P(E_j, V_j, N_k, E_k, V_k, N_k) = \frac{N_T!}{N_j! N_k!} \frac{\iiint \delta(E_j - H_j) \delta(E_k - H_k) dq_j dp_j dq_k dp_k}{\iint \delta(E_T - H_T) dq dp} \quad (6.13)$$

Similar to previous treatments of entropy, we can rewrite the probability in terms of a function<sup>2</sup>:

$$\Omega(E, V, N) = \frac{1}{h^{3N} N!} \iint \delta(E - H) dq dp \quad (6.14)$$

<sup>1</sup>For a plasma, where long range interactions can occur, this is no longer true and different considerations need to be made

<sup>2</sup>The factor of  $h$  comes from quantum mechanical considerations

Thus,

$$P = \frac{\Omega_j \Omega_k}{\Omega_T} \quad (6.15)$$

and

$$S = k \ln \left[ \frac{1}{h^{3N} N!} \iint \delta(E - H) dq dp \right] \quad (6.16)$$

### Second Law of Thermodynamics

We know that the equilibrium condition is what configuration maximises the entropy of the total system. If the entropy is maximised with a constraint, it then follows that if a constraint is applied (think lagrange multiplier) then the entropy must be equal or lower in value. Thus, the total entropy will not decrease when the constraints are removed—this is the second law of thermodynamics.

#### 6.2.1 Zeroeth Law of Thermodynamics

An immediate consequence of the equilibrium conditions is that if two systems are in thermodynamic equilibrium with a third, then they are in equilibrium with each other. This follows from the transitive property of equality:

$$\frac{\partial S_j}{\partial E_j} = \frac{\partial S_l}{\partial E_l} \quad \text{and} \quad \frac{\partial S_k}{\partial E_k} = \frac{\partial S_l}{\partial E_l} \quad \implies \quad \frac{\partial S_j}{\partial E_j} = \frac{\partial S_k}{\partial E_k} \quad (6.17)$$

# Chapter 7

## Thermodynamic Variables

From the equilibrium conditions, we can see that the derivatives of entropy are important thermodynamic quantities characteristic of the equilibrium. The values of these derivatives are related to extensive physical properties of the system, namely temperature, pressure, and chemical potential.

### 7.1 Pressure

#### 7.1.1 Velocity Distribution

We can define the probability density of all locations and momenta as:

$$P(q, p) = \frac{1}{\Omega} \frac{1}{h^{3N} N!} \delta(E - H) \quad (7.1)$$

with

$$\Omega = \frac{1}{h^{3N} N!} \iint \delta(E - H) dq dp \quad (7.2)$$

We are interested in the marginal probability for one particle, with no interactions:

$$P(r_1, p_1) = \frac{\Omega(E - p_1^2/2, V, N - 1)}{N h^3 \Omega(E, V, N)} \quad (7.3)$$

Further integrating over the position,

$$P(p_1) = \frac{V}{N h^3} \frac{\Omega(E - p_1^2/2m, V, N - 1)}{\Omega(E, V, N)} \quad (7.4)$$

Taking the logarithm, we can treat the energy of the single particle as very small compared to the total energy. Thus,

$$\ln P \approx \ln \Omega(E, V, N - 1) - \ln \Omega(E, V, N) - \frac{p_1^2}{2m} \frac{\partial E}{\partial \ln} \Omega(E, V, N - 1) + \ln(V/N h^3)$$

$$\begin{aligned}
&\approx \frac{p_1^2}{2m} \frac{\partial E}{\partial \ln} \Omega(E, V, N) + \text{constants} \\
&\equiv -\frac{p^2}{2m} \beta + \text{constants}
\end{aligned}$$

Where the approximation  $\ln \Omega(E, V, N) - \ln \Omega(E, V, N-1) \sim 1/N$  is taken, and the quantity

$$\beta \equiv \frac{\partial E}{\partial \ln} \Omega(E, V, N) \quad (7.5)$$

Thus, exponentiating the logarithm, the probability becomes:

$$P(p_1) = \left( \frac{\beta}{2\pi m} \right)^{3/2} \exp \left[ -\beta \frac{p_1^2}{2m} \right] \quad (7.6)$$

Separating component wise,

$$P(p_{1,x}) = \left( \frac{\beta}{2\pi m} \right)^{1/2} \exp \left[ -\beta \frac{p_{1,x}^2}{2m} \right] \quad (7.7)$$

Converting to speed,

$$P(v) = \left( \frac{\beta m}{2\pi} \right)^{3/2} 4\pi v^2 \exp \left[ \frac{1}{2} \beta m v^2 \right] \quad (7.8)$$

with

$$\langle v \rangle = \sqrt{\frac{8}{\pi \beta m}} \quad v_{mp} = \sqrt{\frac{2}{\beta m}} \quad \langle v^2 \rangle = \frac{3}{\beta m} \quad (7.9)$$

### 7.1.2 Temperature and the Ideal Gas Law

Consider a gas in a box colliding with a small, flat portion of the wall with area  $A$ . When a particle collides with a wall, the momentum transferred is in the direction of the normal to a surface. During a small time  $\Delta t$ , only particles with  $v_x > 0$  and within  $v_x \Delta t$  will hit the wall.

The average number of particles that will strike the section of wall will be

$$dN_c(p_x, \Delta t) = \underbrace{\left( \frac{N}{V} \right)}_{\text{density}} \cdot \underbrace{(P(p_x) dp_x)}_{\text{prob.}} \cdot \underbrace{\left( \frac{A p_x}{m} \right) \Delta t}_{\text{volume}} \quad (7.10)$$

The impulse of the collisions is  $\Delta p_x = 2p_x$ . Additionally, we are only concerned about pressure, or force per area. The pressure is then:

$$P = \int \Delta p_x dN_c \quad (7.11)$$

Substituting and evaluating this integral, we get

$$PV = N\beta^{-1} \quad (7.12)$$

Comparing to the empirical ideal gas law,

$$PV = Nk_B T = nRT \quad (7.13)$$

We see that the inverse temperature, thermodynamic beta, is

$$\beta = \frac{1}{k_B T} \quad (7.14)$$

Thus, we see that the equilibrium condition for energy exchange is when the two systems have the same temperature  $T$ . Further, from the ideal gas law, and the fact that  $P$ ,  $V$ ,  $N$  are all positive, we see that temperature must be positive. Thus, an absolute temperature scale must be used—the one we use is Kelvin. Given the Kelvin scale and the ideal gas law, the Boltzmann constant is experimentally determined to be

$$k_B = 1.38 \times 10^{-23} \text{ J} \cdot \text{K}^{-1} \quad (7.15)$$

## 7.2 Derivatives of Entropy

Taking the derivative of entropy with respect to temperature, we get

$$\left( \frac{\partial S}{\partial V} \right)_{E,N} = \frac{P}{T} \quad (7.16)$$

Taking the derivative of entropy wrt energy yields:

$$\left( \frac{\partial S}{\partial E} \right)_{V,N} = \frac{1}{T} \quad (7.17)$$

Evaluating this derivative with the expression for entropy, we obtain

$$\frac{3k_B N}{2E} = \frac{1}{T}$$

Rearranging, we get the a case of the equipartition theorem:

$$\frac{E}{N} = 3 \cdot \frac{1}{2} k_B T \quad (7.18)$$

The chemical potential is defined to be

$$-\frac{\mu}{T} = \left( \frac{\partial S}{\partial N} \right)_{E,V} \quad (7.19)$$

### 7.3 Asymmetric Pistons

Imagine two boxes connected by a movable piston. The piston for box  $j$  has area  $A_j$ , and the piston for box  $k$  has area  $A_k$ . As the piston moves, the total volume of the system changes. At equilibrium, the force on the piston on each side is given  $F = P_j A_j = P_k A_k$ , and the temperatures are equal. Thus, we have

$$\frac{F}{T} = \frac{P_i A_i}{T_i} = A_i \left( \frac{\partial S_i}{\partial V_i} \right)_{E_i, N_i}$$

Thus, the equilibrium condition is actually

$$A_j \frac{\partial S_j}{\partial V_j} = A_k \frac{\partial S_k}{\partial V_k} \quad (7.20)$$

Note that this implies that the partial derivatives are not equal.

### 7.4 Differential form of thermodynamics

Thermodynamics is concerned primarily with small changes in macroscopic parameters. Using the total differential of entropy,

$$dS = \frac{\partial S}{\partial E} dE + \frac{\partial S}{\partial V} dV + \frac{\partial S}{\partial N} dN$$

and substituting in the equations of state, we obtain the differential form of thermodynamics in the entropy representation:

$$dS = \frac{1}{T} dE + \frac{P}{T} dV - \frac{\mu}{T} dN \quad (7.21)$$

This is valid for all thermodynamic systems.

#### 1st Law

The quantity  $E$  is most often denoted  $U$ . Notice, by multiplying Equation 7.21 with temperature, we obtain

$$T dS = dU + P dV - \mu dN$$

$$dU = T dS - P dV + \mu dN$$

This last equation is a differential statement of the First Law of Thermodynamics.

### 7.5 Measurements and Reservoirs

A measuring device is a tiny device where when it comes into equilibrium with a system, the system barely changes. A reservoir is the opposite; when a system get into equilibrium with a reservoir, the reservoir barely changes.



## Part II

# Postulates of Thermodynamics



# Chapter 8

## Postulates and Laws of Thermodynamics

Statistical mechanics and thermodynamics are different ways to approach thermal physics. The same equations can be derived from either theory, but they are not necessarily with the same ease. The different origins lead to different assumptions between the theories.

Statistical mechanics assumes that the particles exist and obey laws of classical and quantum mechanics. The statistics on these particles give justification for the laws of thermodynamics. Thermodynamics on the other hand, was developed from more experimental evidence, and used postulates to derive behaviours.

Statistical mechanics deals with the exact state of a system, or an exact point in phase space. In contrast, thermodynamics deals with the macroscopic state, which can be characterised by a handful of variables. This macroscopic state can correspond to a collection of different microscopic states, but it cannot be used to determine the precise microscopic state.

An equilibrium state is a special kind of a macroscopic state; it does not change with time. This equilibrium state can be totally characterized by a small number of variables. For example, the ideal gas law is one way to characterize the equilibrium state for an ideal gas.

A *state function* is a function of a small number of parameters that describes an equilibrium state. One of the most important state functions, which we have already discussed, is *entropy*.

### 8.1 Postulates

The postulates of thermodynamics that are used in Swendsen are:

1. There exist equilibrium states that are characterized uniquely by a small number of extensive variables.
2. There exists a state function (entropy) such that equilibrium states maximise this state function.
3. The total entropy of two systems is the sum of the individual entropies.

4. Entropy is a continuous and differentiable
5. (Optional) Entropy is an extensive property of a system.

$$S(\lambda E, \lambda V, \lambda N) = \lambda S(E, V, N)$$

6. (Optional) Entropy is a monotonically increasing function of energy for equilibrium values of energy.
7. (Optional) Entropy is always non-negative

Note that postulate #2 implies the concept of irreversibility, and consequently the arrow of time. Additionally, note that postulate #3 ignores interactions between particles in different subsystem; the hamiltonian includes a coupling term between the boxes. This is postulate is an approximation, albeit a good one, as interactions are over very short lengths.

The optional postulates are those which are usually true, but there still exist some systems where they do not hold.

## 8.2 Laws

0. Two systems in equilibrium with a third are in equilibrium with each other.
1. Heat is a form of energy transfer, and energy is always conserved.
2. After the release of a constraint on a system, the entropy can never decrease.
3. The entropy of a system approaches a constant as temperature approaches zero.

# Chapter 9

## Perturbations of State Functions

State functions do not depend on how systems got into the present state, only the extensive properties of the state at that time. However, much of thermodynamics is concerned with small changes, and larger changes can be obtained by summing/integrating over these small changes.

### 9.1 Energy Conservation

Previously, we denoted the energy of a system to be  $E$ . However, hereon, we shall denote the internal energy of a system to instead be  $U$ .

The first law of thermodynamics states that heat and work are changes in the internal energy of a system. This can be written in differential form as

$$dU = dQ + dW \quad (9.1)$$

where  $Q$  and  $W$  denote heat and work respectively. The sign convention being used is when energy is added to a system, the sign is positive; if energy is being used, the sign is negative.

Note the slightly different notation of the differentials on each side. The regular differential on the LHS denotes an *exact* differential, while the slashed differential denotes an *inexact* differential.

#### Differentials

What is an inexact differential? Consider the differential below:

$$dF = f_1(x, y) dx + f_2(x, y) dy$$

In general, the integral of this differential depends on the path  $(x(t), y(t))$  that is being taken to integrate it. Using this parametrization, we can rewrite the differential as:

$$dF = f_1(t) \frac{dx}{dt} dt + f_2(t) \frac{dy}{dt} dt$$

Or, more compactly,

$$dF = [f \circ \gamma(t)] \cdot [\gamma'(t)] dt \quad (9.2)$$

if, however, a differential can be written in the form

$$dF = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy$$

it is an exact differential. This can be verified using Clairaut's theorem. If

$$\frac{\partial f_1}{\partial y} = \frac{\partial^2 F}{\partial x \partial y} = \frac{\partial f_2}{\partial x}$$

on a simply connected region, then the differential form is exact. For 1-forms, which are most likely to be used, an exact differential form is given:

$$dF = \sum_i \frac{\partial F}{\partial x^i} dx^i = \nabla F \cdot dr$$

Once again, Clairaut's theorem can be used to verify if a differential is exact.

An integrating factor can be multiplied to an inexact differential form to yield an exact differential form.

$$dG = r dF \quad (9.3)$$

In 2 dimensions, this integrating factor can be found by solving

$$\frac{\partial}{\partial y}(rf_1) = \frac{\partial}{\partial x}(rf_2)$$

Note that state functions have exact differentials.

## 9.2 Integrating Factors in Classical Ideal Gas

A gas does work on a piston with area  $A$ . The work being done is given

$$dW = -F dx$$

However, we can write

$$F = PA$$

substituting in, we have

$$dW = -P dV \quad (9.4)$$

Clearly, we have an integrating factor of  $1/P$ , as it is what separates this inexact differential from being the exact differential  $dV$ .

Similarly, for an increase of entropy for a small amount of heat  $dQ$ , is

$$dS = S(U + dQ, V, N) - S(U, V, N) = \frac{\partial S}{\partial U} dQ$$

Or,

$$dS = \frac{1}{T} dQ \quad (9.5)$$

Thus,  $\frac{1}{T}$  is an integrating factor from work to entropy.

## 9.3 Work and Heat

Work is defined to be

$$\int_A^B F \cdot dr \quad (9.6)$$

and can be defined for a variety of different forces. For example, spring tension force has work

$$dW = \sigma dL$$

and battery emf has work

$$dW = \mathcal{E} dq$$

Work is an extensive quantity.

Heat has 3 forms of transfer. They are conduction, convection, and radiation. Conduction is due to microscopic collisions on an interface. Convection is due to the transport of molecules through a fluid. Radiation is due to the emission and absorption of electromagnetic radiation.

Using the First Law of Thermodynamics, we can write

$$dU = T dS - P dV \quad (9.7)$$

Note that this is only valid if there is no exchange of particles. In a previous chapter, derived in an aside, we showed the more general expression

$$dU = T dS - P dV + \mu dN \quad (9.8)$$





# Chapter 10

## Heat Engines and Refrigerators

A heat engine is a thermodynamic cycle that converts heat into useable work. Quasi-static processes are useful to study these cycles.

A quasi-static process is a process that happens slowly enough where the system always remains in equilibrium; the entropy does not increase, and so the process is *reversible*. While such a process does not truly exist, it is a useful construction. All real thermodynamic processes result in the increase of entropy, and thus are irreversible.

### 10.1 Heat Capacity

The heat capacity of an object is defined as

$$C = \frac{dQ}{dT} \quad (10.1)$$

Often, we are interested in constant volume and constant pressure processes. Respectively,  $C_V$  and  $C_P$  denote heat capacity at constant volume and constant pressure.

Additionally, we can make the heat capacity into the intensive property *specific heat capacity*  $c$ , which is the heat capacity per unit mass.

Under the assumption that particles are not exchanged, under constant volume  $dV = 0$ , we get the heat capacity to be

$$C_V = \frac{dQ}{dT} = \frac{dU}{dT} \quad (10.2)$$

Using the energy for an ideal monatomic gas,<sup>1</sup>

$$U = \frac{3}{2}Nk_B T$$

---

<sup>1</sup>The reason that we specify a monatomic gas is due to the equipartition theorem. Each quadratic degree of freedom has  $\frac{1}{2}k_B T$  worth of energy; an ideal diatomic gas would have energy  $U = \frac{5}{2}k_B T$ , as there are 5 quadratic degrees of freedom (3 translation and 2 rotation)

we have

$$C_V = \frac{3}{2}k_B N \quad (10.3)$$

For constant pressure, we can use the ideal gas law

$$PV = Nk_B T$$

to write

$$C_P = \frac{dU + P dV}{dT} = \frac{dU + Nk_B dT}{dT} = C_V + Nk_B \quad (10.4)$$

This can be rearranged to give Mayer's relation for an ideal gas:

$$nR = C_P - C_V \quad (10.5)$$

Thus, we must have

$$C_P > C_V \quad (10.6)$$

Further, we can compute

$$\gamma \equiv \frac{C_P}{C_V} = \frac{C_V + Nk_B}{C_V} = \frac{5}{3} = \frac{f+2}{f} \quad (10.7)$$

This quantity, known as the *adiabatic index*, plays an important role in adiabatic processes.

## 10.2 Adiabatic Processes

In an adiabatic process, no heat is transferred. In an adiabatic expansion or compression, the volume of the gas is changed very rapidly, as to not allow any time for heat to transfer.

Differentiating the ideal gas law, we then have:

$$P dV + V dP = nR dT = (C_P - C_V) dT$$

Because there is no heat, we can write

$$C_V = \frac{dU}{dT} = -\frac{P dV}{dT}$$

Plugging this in, we get:

$$P dV + V dP = (\gamma - 1)C_V dT = (1 - \gamma)P dV$$

moving terms and simplifying,

$$\gamma \frac{dV}{V} + \frac{dP}{P} = 0$$

Integrating, we get

$$P_1 V_1^\gamma = P_2 V_2^\gamma \quad (10.8)$$

Plugging in the ideal gas law, we can write the effect on temperature as:

$$T_1 V_1^{\gamma-1} = T_2 V_2^{\gamma-1} \quad (10.9)$$

Thus, as volume decreases, the temperature increases. This is because work is being done on the gas to compress it. In contrast, with adiabatic expansion, the gas does work and the temperature must decrease.

Adiabatic expansion can be contrasted with *free expansion*, where a gas is allowed to expand into a vacuum. Because there is no pressure, the work a gas does is zero, and the temperature does not change.

## 10.3 Thermodynamic Cycles

Thermodynamic cycles operate on two principles. Because a thermodynamic cycle is cyclic, the entropy change after a complete cycle (assuming quazi-static/ideal processes) must be zero (although it can be non-zero during the cycle).

$$dS = \frac{dQ_H}{T_H} + \frac{dQ_L}{T_L} = 0 \quad (10.10)$$

Further, the work is the amount of energy extracted from the heat:

$$dW = dQ_H + dQ_L \quad (10.11)$$

### 10.3.1 Heat Engine

A heat engine is a thermodynamic cycle that converts (useable) heat into work. The prototypical heat engine is the Carnot Cycle. From the hot reservoir, we have  $dQ_H > 0$  coming into the engine, and  $-dQ_L > 0$  removed from the system.

The ideal efficiency of a heat engine is the Carnot efficiency, and is defined as the ratio of the work extracted from the total amount of heat entering the engine:

$$\eta \equiv \frac{dW}{dQ_H} = 1 - \frac{T_L}{T_H} \quad (10.12)$$

The first law tells us that  $dW \leq dQ_H$  (you can't win), and the second law says that  $dS > 0$  (you can't break even).

### 10.3.2 Refrigerator

By reversing a heat engine, we can use mechanical work to pump heat from a cold reservoir to a hot reservoir. The coefficient of performance for a refrigerator is defined to be the

amount of heat extracted from the cold reservoir divided by the amount of work required to do so:

$$\epsilon_R = \frac{dQ_L}{-dW} = \frac{T_L}{T_H - T_L} \quad (10.13)$$

In a refrigerator, a compressor compresses a refrigerant vapour, increasing the temperature above ambient. The ambient air cools the compressed vapour until it condenses in the condenser. Then, the liquid cools as it passes through an expansion valve and enters the refrigerator. The liquid refrigerant absorbs heat from the refrigerator to evaporate, and the gaseous refrigerant re-enters the compressor.

### 10.3.3 Heat Pump

A heat pump operates in a similar way, except it uses the heat outside a system to heat a system. Its coefficient of performance is defined to the amount of heat pumped into the hot reservoir divided by the amount of work required to do so:<sup>2</sup>

$$\epsilon_{HP} = \frac{-dQ_H}{dW} = \frac{T_H}{T_H - T_L} \quad (10.14)$$

Interestingly, a heat pump often has a much higher efficiency than an electrical resistive heating element.

## 10.4 Carnot Cycle

The Carnot Cycle is a four-stage process:

1. Isothermal expansion in contact with a hot reservoir
2. Adiabatic expansion
3. Isothermal compression in contact with cold reservoir
4. Adiabatic compression

### Isothermal Expansion

Because the change in temperature is zero, the change in  $U = \frac{3}{2}nRT$  is also zero, thus,

$$W_{12} = \int_{V_A}^{V_B} P dV = Nk_B T \int_{V_A}^{V_B} dV V^{-1} = Nk_B T \ln \left( \frac{V_B}{V_A} \right) = Q_H$$

Further, the change in entropy is given:

$$\Delta S_{12} = \frac{Q_H}{T_H}$$

---

<sup>2</sup>Note that while the difference between a refrigerator and heat pump may seem superficial at first, the difference between the two cycles is the reservoir that we care about.

### Adiabatic Expansion

The ideal gas is thermally isolated and allowed to expand adiabatically. Thus, the temperature decreases to  $T_L$ . Using the equation for adiabatic expansion (and plugging in  $\gamma = 5/3$ )

$$T_H V_2^{2/3} = T_L V_3^{2/3} \implies V_3 = V_2 \left( \frac{T_H}{T_L} \right)^{3/2}$$

The work being done is more complicated to calculate and has been temporarily omitted. Because no heat is transferred, the change in entropy is

$$\Delta S_{23} = 0$$

### Isothermal Compression

Once again, temperature is held constant while the gas is expanded. Similar to the isothermal expansion, we can find the work done as

$$W_{34} = N k_B T_L \ln \left( \frac{V_4}{V_3} \right) = -Q_L$$

and the entropy is

$$\Delta S_{34} = -\frac{Q_L}{T_L}$$

### Adiabatic Compression

Again, the volume and temperature obey the relation

$$T_L V_4^{2/3} = T_H V_1^{2/3}$$

and the change in entropy is

$$\Delta S_{41} = 0$$

## 10.4.1 Total Cycle

Over a complete cycle, the work done by the gas is given

$$W = W_{12} + W_{23} - W_{34} - W_{41} \tag{10.15}$$

and the change in internal energy tells us

$$0 = \Delta U = \Delta Q - W \implies W = Q_H - Q_L$$

The total change in entropy is also zero:

$$0 = \Delta S = S_{12} + S_{23} + S_{34} + S_{41} = \frac{Q_H}{T_H} - \frac{Q_L}{T_L}$$

Thus, we have

$$\frac{Q_H}{Q_L} = \frac{T_H}{T_L} \quad (10.16)$$

and the volumes are related by

$$\frac{T_H}{T_L} = \frac{Q_H}{Q_L} = \frac{Nk_B T_H \ln\left(\frac{V_2}{V_1}\right)}{Nk_B T_L \ln\left(\frac{V_3}{V_4}\right)}$$

$$\frac{V_2}{V_1} = \frac{V_3}{V_4} \quad (10.17)$$

This cycle can be visualized in the PV diagram in Figure 10.1 and the TS diagram in Figure 10.2

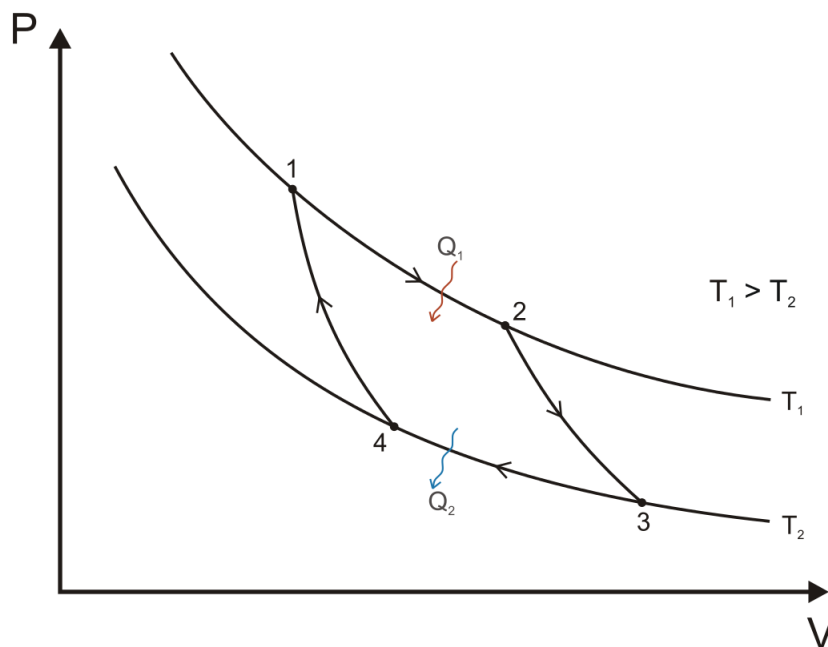


Figure 10.1: PV diagram of Carnot Cycle

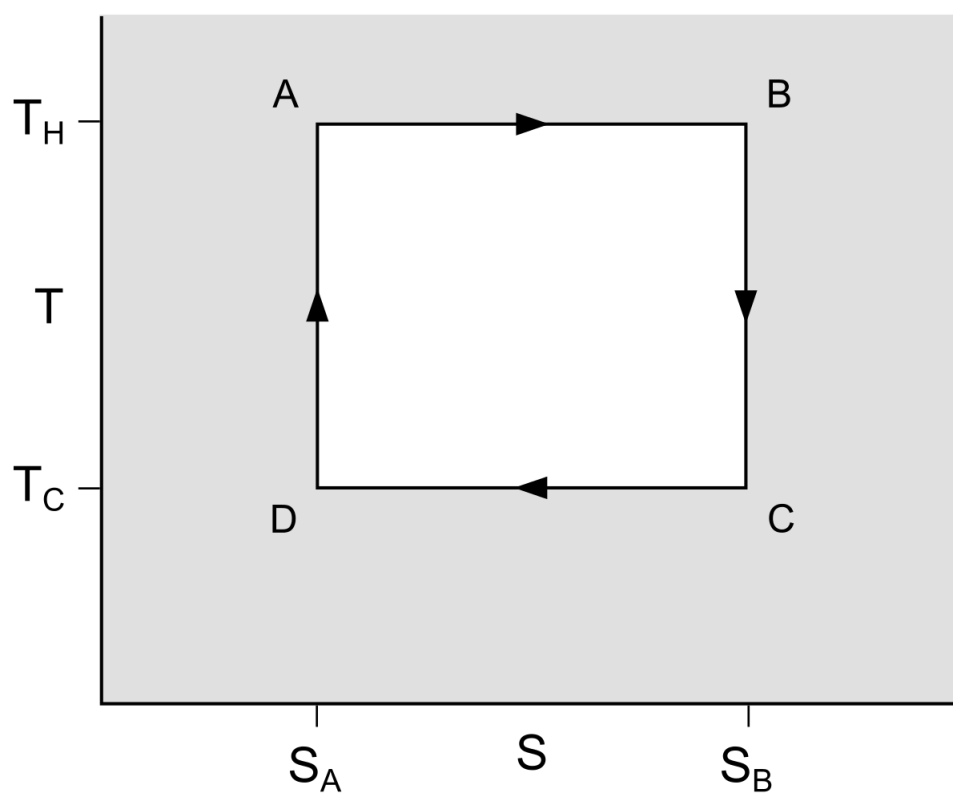


Figure 10.2: TS diagram of Carnot Cycle

Negative Temperature

Recall that temperature is defined by

$$k_B\beta = \frac{1}{T} = \frac{\partial S}{\partial E}$$

If entropy decreases with an increase in energy, then the temperature is *negative*. If a system with negative temperature is put in thermal contact with a system with positive temperature, the equilibrium point is not zero, but rather infinite! Rather than colder than absolute zero, *a negative temperature is hotter than infinite kelvin*.

The discontinuity of traditional temperature motivates the importance of thermodynamic beta, as absolute zero is  $+\infty$ , the value at 0 is continuous.



# Chapter 11

## Thermodynamic Potentials

The state functions  $S(U, V, N)$  and  $U(S, V, N)$  contain all the thermodynamic information for a system at equilibrium. It is not always easy to extract information from these quantities. Many processes are done at constant temperature; chemistry is done at constant pressure. Thus, it is often more convenient to express thermodynamic relations in terms of intensive properties  $T$  and  $P$  which can be more easily measured than the extensive  $S$ ,  $V$ , and  $N$ . Such relations can be obtained using the *thermodynamic potentials* by taking Legendre transformations.

### 11.0.1 Legendre Transformation

The Legendre transform is defined as

$$\tilde{f}(p) = f^*(p) \equiv px(p) - f(x(p)) \quad p \equiv \frac{\partial f}{\partial x} \quad (11.1)$$

#### Legendre Transform Example

The Legendre transform of the function

$$f(x) = 2 - \ln(x - x_0)$$

can be found:

$$p = \frac{\partial f}{\partial x} = \frac{1}{x - x_0} \implies x = x_0 + \frac{1}{p}$$
$$\tilde{f} = x_0 p - 1 + \ln(p)$$

If we consider the tangent line to a function  $f$  at a point  $x_0$ ,

$$y = f(x_0) + \frac{\partial f}{\partial x}(x - x_0)$$

$$\begin{aligned} y &= f(x_0) + p(x - x_0) \\ y(0) &= f(x_0) - px_0 = -\tilde{f} \end{aligned} \quad (11.2)$$

or, the Legendre transform maps a function to the negative of the y intercept of the tangent line at the point.

In particular, the Legendre transformation is useful for concave and convex functions. A convex function is one where the segment connecting two points on the function lies completely above the function; a concave function is the negative of a convex function. In 1D, a concave function is defined by  $f'' \leq 0$ , and a convex function  $f'' \geq 0$ .

The most famous Legendre transform is the transformation from a Lagrangian to a Hamiltonian.

## 11.1 Transforms of Energy

### 11.1.1 Helmholtz Free Energy

Recall that temperature is defined

$$\frac{1}{T} = \frac{\partial S}{\partial U}$$

Thus,

$$T = \frac{\partial U}{\partial S} \quad (11.3)$$

Taking the Legendre Transform with respect to temperature  $F = U[T]$

$$\begin{aligned} f(S) &= U(S) \\ \tilde{f}(T) &= TS - U \end{aligned}$$

We define the quantity  $\tilde{f}(T) = -F$ , so

$$F = U - TS$$

The differential of this quantity is

$$dF = dU - S dT - T dS$$

Substituting the differential form of thermodynamics Equation 9.8, we get

$$dF = -S dT - P dV + \mu dN \quad (11.4)$$

so,

$$F(T, V, N) = U - TS \quad (11.5)$$

Thus, we have the equations of state

$$S = -\frac{\partial F}{\partial T} \quad (11.6a)$$

$$P = -\frac{\partial F}{\partial V} \quad (11.6b)$$

$$\mu = \frac{\partial F}{\partial N} \quad (11.6c)$$

### 11.1.2 Enthalpy

Enthalpy is the Legendre transform with respect to pressure. From Equation 9.8, we see that

$$P = -\frac{\partial U}{\partial V}$$

Thus, the Legendre transform  $H = U[P]$  is:

$$f(V) = U(V)$$

$$\tilde{f}(P) = -PV - U$$

Defining the quantity  $\tilde{f}(V) \equiv -H$ ,

$$H = U + PV$$

Taking the total differential,

$$dH = dU + P dV + V dP$$

and substituting in Equation 9.8

$$dH = T dS + V dP + \mu dN \quad (11.7)$$

so

$$H(S, P, N) = U + PV \quad (11.8)$$

One of the equations of state is given

$$V = \frac{\partial H}{\partial P} \quad (11.9)$$

The enthalpy is used most for chemical reactions. Most chemical reactions are done at constant pressure and constant number<sup>1</sup>, and so

$$dH = T dS = dQ$$

---

<sup>1</sup>gross oversimplification

### 11.1.3 Gibbs Free Energy

The gibbs free energy is a 2D legendre transform of energy with respect to both temperature and pressure. The gibbs free energy is given:

$$G(T, P, N) = U - TS + PV \quad (11.10)$$

and has the differential form

$$dG = -S dT + V dP + \mu dN \quad (11.11)$$

The equations of state are

$$S = \frac{\partial G}{\partial T} \quad (11.12a)$$

$$V = \frac{\partial G}{\partial P} \quad (11.12b)$$

#### Other potentials

With three variables  $S, V, N$  defining  $U$ , we can take a total of  $2^3 = 8$  thermodynamic potentials.

We could also start with the state function of entropy, defining  $\tilde{S} = S/k_B$

$$d\tilde{S} = \beta dU + \beta P dV - \beta \mu dN$$

The legendre transforms with respect to  $\beta$ ,  $(\beta P)$ , and  $(\beta \mu)$  are known as the Massieu Functions

# Chapter 12

## Consequences of Extensivity

Recall that entropy is extensive if it satisfies:

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

that is,  $S, U, V, N$  are proportional to the size of the system.

Entropy is considered additive if

$$S_{A,B}(U_A, V_A, N_A; U_B, V_B, N_B) = S_A + S_B$$

Note that both of these properties can be violated if there are surface effects, or if molecular interactions are short range.

Extensivity implies additivity, but not vice versa.

### 12.1 Euler Equation

Energy is also an extensive property. If we take the derivative wrt the coefficient  $\lambda$ , we get

$$U(S, V, N) = \sum_i \frac{\partial U(\lambda x_1, \lambda x_2, \lambda x_3)}{\partial \lambda x_i} \frac{\partial \lambda x_i}{\partial \lambda} = \sum_i \frac{\partial U(\lambda \vec{x})}{\partial \lambda} x_i$$

Fixing  $\lambda = 1$ ,

$$U = S \frac{\partial U}{\partial S} + V \frac{\partial U}{\partial V} + N \frac{\partial U}{\partial N}$$

Substituting,

$$U = TS - PV + \mu N \tag{12.1}$$

or,

$$S = \frac{U}{T} + \frac{PV}{T} - \frac{\mu N}{T}$$

## 12.2 Gibbs-Duhem Relation

If a system is extensive, the intensive parameters are not independent. Taking the differential of the Euler equation,

$$dU = T dS + S dT - P dV - V dP + \mu dN + N d\mu$$

we can subtract the first law to obtain the Gibbs-Duhem Relation:

$$0 = S dT - V dP + N d\mu \quad (12.2)$$

Generally, we consider  $P, T$  to be the free parameters:

$$d\mu = -\frac{S}{N} dT + \frac{V}{N} dP \quad (12.3)$$

If we use a more general formulation of the first law allowing for different species,

$$dU = T dS - P dV + \sum_j^r \mu_j dN_j$$

we see that the relation becomes

$$0 = S dT - V dP + \sum_j^r N_j d\mu_j$$

so there is the loss of one free parameter.

Similarly, we can rewrite the Gibbs-Duhem relation in terms of the Euler equation for entropy:

$$d\left(\frac{\mu}{T}\right) = \frac{U}{N} d\left(\frac{1}{T}\right) + \frac{V}{N} d\left(\frac{P}{T}\right) \quad (12.4)$$

## 12.3 Application to Ideal Gas

Taking the ideal gas law and equipartition theorem for a monatomic gas, we can rewrite in terms of intensive variables  $u = U/N$  and  $v = V/N$

$$\frac{P}{T} = k_B v^{-1} \quad \frac{1}{T} = \frac{3}{2} k_B u^{-1}$$

Plugging into the second form of the Gibbs-Duhem relation,

$$d\left(\frac{\mu}{T}\right) = -\frac{3}{2} k_B u^{-1} du - k_B v^{-1} dv$$

we can integrate and obtain a third equation of state:

$$\frac{\mu}{T} = -k_B \left[ \frac{3}{2} \ln \frac{U}{N} + \ln \frac{V}{N} + X \right] \quad (12.5)$$

Plugging the euler equation into the thermodynamic potentials, we obtain

$$F = -PV + \mu N \quad (12.6a)$$

$$H = TS + \mu N \quad (12.6b)$$

$$G = \mu N \quad (12.6c)$$

Recall that this applies only to extensive systems.

### Legendre Transforms

Notice that legendre transforms exchange a differential in terms of  $A \, dB$  for  $-B \, dA$ . This corresponds to the subtraction of the term  $AB$  in the overall function. This is because:

$$d(AB) = A \, dB + B \, dA$$

subtracting this differential shows how the differential gets exchanged; the legendre transform swaps the variables in the differential term.





# Chapter 13

## Thermodynamic Identities

### 13.1 Second Derivatives

Recall the first derivatives of state functions give the properties of a thermodynamic system. The second derivatives yield the way these parameters change. However, many second derivatives can be rewritten in terms of other combinations.

Take for example,  $S(U, V, N)$ , Naively, there are 6 second derivatives, but often  $N$  is held fixed, reducing the number of second derivatives to 3. All other second derivatives can be written in terms of these three second derivatives, in relations known as *thermodynamic identities*

The three standard second derivatives are chosen by ease of measurement. For this class, we choose the *coefficient of thermal expansion*

$$\alpha = \frac{1}{V} \left( \frac{\partial V}{\partial T} \right)_{P,N} \quad (13.1)$$

*isothermal compressibility*

$$\kappa_T = -\frac{1}{V} \left( \frac{\partial V}{\partial P} \right)_{T,N} \quad (13.2)$$

and specific heat (per particle)

$$c_P = \frac{T}{N} \left( \frac{\partial S}{\partial T} \right)_{P,N} \quad \text{or} \quad c_V = \frac{T}{N} \left( \frac{\partial S}{\partial T} \right)_{V,N} \quad (13.3)$$

Usually,  $c_P$  is easier to measure for solids, but  $c_V$  for gases.

### 13.2 Maxwell Relations

Recall that  $dU, dS, dF, dH, dG, \dots$  are exact differentials. Fixing  $dN = 0$ , we have

$$dU = T dS - P dV$$

This implies

$$\left(\frac{\partial T}{\partial V}\right)_{S,N} = -\left(\frac{\partial P}{\partial S}\right)_{T,N} \quad (13.4)$$

Similarly, for the Helmholtz free energy,

$$\begin{aligned} dF &= -S dT - P dV \\ -\left(\frac{\partial S}{\partial V}\right)_{T,N} &= -\left(\frac{\partial P}{\partial T}\right)_{V,N} \end{aligned} \quad (13.5)$$

Recall that each term in a thermodynamic potential is given as  $A dB$  or  $-B dA$ , related by legendre transforms. Given there are 3 such terms, there are 8 such terms. These are

$$T dS \quad \text{or} \quad -S dT \quad (13.6a)$$

$$V dP \quad \text{or} \quad -P dV \quad (13.6b)$$

$$\mu dN \quad \text{or} \quad -N d\mu \quad (13.6c)$$

For example, the derivative

$$\left(\frac{\partial T}{\partial P}\right)_{S,\mu}$$

comes from the differential

$$+T dS + V dP - N d\mu$$

because the terms  $S, \mu$  are held constant, they correspond to the differentials. Because  $T$  is being differentiated. Because the temperature is being differentiated wrt  $P$ , there must also be a differential  $dP$  term. Fixing  $d\mu = 0$ , we obtain the corresponding derivative:

$$\left(\frac{\partial T}{\partial P}\right)_{S,\mu} = \left(\frac{\partial V}{\partial S}\right)_{P,\mu} \quad (13.7)$$

If, however, a corresponding potential does not exist, the relation should be found for the reciprocal, then take the reciprocal of the inverse. For example,

$$\left(\frac{\partial P}{\partial T}\right)_{S,N}$$

would correspond to a  $-S dT$  term (following from the denominator), but  $S$  being held constant implies a term  $T dS$ , which is contradictory. Instead we consider the reciprocal of the derivative

$$\left(\frac{\partial T}{\partial P}\right)_{S,N}$$

corresponds to the potential

$$T dS + V dP + \mu dN$$

So, the reciprocal relation is

$$\left(\frac{\partial T}{\partial P}\right)_{S,N} = \left(\frac{\partial V}{\partial S}\right)_{T,N}$$

which yields the desired relation

$$\left(\frac{\partial P}{\partial T}\right)_{S,N} = \left(\frac{\partial S}{\partial V}\right)_{T,N} \quad (13.8)$$

### 13.3 “Insert $\partial(P, T)$ ”—Jacobians

The Jacobian is another useful tool for manipulating partial derivatives. Recall that a Jacobian is defined as:

$$\frac{\partial(\vec{u})}{\partial(\vec{v})} = \begin{vmatrix} \frac{\partial u_1}{\partial v_1} & \frac{\partial u_1}{\partial v_2} & \cdots \\ \frac{\partial u_2}{\partial v_1} & \frac{\partial u_2}{\partial v_2} & \cdots \\ \vdots & \vdots & \ddots \end{vmatrix} \quad (13.9)$$

Recall that the determinant is an alternating map; thus, switching the order of variables yields the negative determinant:

$$\frac{\partial(u, v)}{\partial(x, y)} = -\frac{\partial(v, u)}{\partial(x, y)} = -\frac{\partial(u, v)}{\partial(y, x)} = \frac{\partial(v, u)}{\partial(y, x)}$$

More generally, for permutations  $\sigma, \tau$ ,

$$\frac{\partial(\sigma\vec{u})}{\partial(\tau\vec{v})} = \frac{\partial(\vec{u})}{\partial(\vec{v})} \frac{\text{sgn}(\sigma)}{\text{sgn}(\tau)} \quad (13.10)$$

If variables occur on the top and bottom of the jacobian, we then have:

$$\frac{\partial(u, y, z)}{\partial(x, y, z)} = \left(\frac{\partial u}{\partial x}\right)_{y,z} \quad (13.11)$$

Further, there is a chain rule for jacobians:

$$\frac{\partial(\vec{u})}{\partial(\vec{v})} = \frac{\partial(\vec{u})}{\partial(\vec{t})} \cdot \frac{\partial(\vec{t})}{\partial(\vec{v})} \quad (13.12)$$

Following immediatly from the chain rule is:

$$\frac{\partial(\vec{a})}{\partial(\vec{c})} \cdot \frac{\partial(\vec{b})}{\partial(\vec{d})} = \frac{\partial(\vec{a})}{\partial(\vec{d})} \cdot \frac{\partial(\vec{b})}{\partial(\vec{c})} = \frac{\partial(\vec{b})}{\partial(\vec{c})} \cdot \frac{\partial(\vec{a})}{\partial(\vec{d})} \frac{\partial(\vec{d})}{\partial(\vec{a})} = \frac{\partial(\vec{b})}{\partial(\vec{c})} \cdot \frac{\partial(\vec{a})}{\partial(\vec{c})} \quad (13.13)$$

The product of a jacobian trivially shows that the reciprocal of a jacobian is simply:

$$1 = \frac{\partial(\vec{u})}{\partial(\vec{v})} \cdot \frac{\partial(\vec{v})}{\partial(\vec{u})} \quad (13.14)$$

Using these tricks with the jacobian we can rewrite derivatives in terms of the standard set. Using Equation 13.11 we see the standard set can be written in terms of

$$\frac{\partial V}{\partial T} = \quad (13.15a)$$

$$\frac{\partial V}{\partial P} = \quad (13.15b)$$

$$\frac{\partial S}{\partial T} = \quad (13.15c)$$

For example, we want to find the equivalent of

$$\left(\frac{\partial P}{\partial T}\right)_V$$

we rewrite as

$$\begin{aligned} \left(\frac{\partial P}{\partial T}\right)_V &= \frac{\partial(P, V)}{\partial(T, V)} = \frac{\partial(P, V)}{\partial(P, T)} \cdot \frac{\partial(P, T)}{\partial(T, V)} = -\frac{\partial(V, P)}{\partial(T, P)} \cdot \frac{1}{\frac{\partial(V, T)}{\partial(P, T)}} \\ &= \frac{1}{V} \left(\frac{\partial V}{\partial T}\right)_P \cdot \frac{1}{\frac{1}{V} \left(\frac{\partial V}{\partial P}\right)_T} \\ &= \alpha \cdot \frac{1}{\kappa_T} \\ &= \frac{\alpha}{\kappa_T} \end{aligned}$$

## 13.4 “Divide Through”

If we have an exact differential

$$dF = \sum_i \left(\frac{\partial F}{\partial x^i}\right)_{x \setminus x^i} dx^i$$

we can “divide through by another quantity to obtain:

$$\left(\frac{\partial F}{\partial u}\right)_{v \setminus u} = \sum_i \left(\frac{\partial F}{\partial x^i}\right)_{x \setminus x^i} \left(\frac{\partial x^i}{\partial u}\right)_{v \setminus u}$$

For example, using

$$dU = T dS - P dV$$

we can divide through by  $dT$  to obtain

$$\left(\frac{dU}{dT}\right)_{P, N} = T \left(\frac{\partial S}{\partial T}\right)_{P, N} - P \left(\frac{\partial V}{\partial T}\right)_{P, N}$$

## 13.5 Joule-Thomson Effect

A gas at a constant, high pressure  $P_A$  is forced through a porous plug into a container with a constant, low pressure  $P_B$ .

Initially, all of the gas is in container  $A$ , so  $U_i = U_A$ , and at the end, all of the gas is in container  $B$ , so  $U_f = U_A$ . Work  $P_A V_A$  is done on gas to evacuate  $A$ , while the gas does work  $P_B V_B$  to fill  $B$ . Thus, the work is given

$$W = P_A V_A - P_B V_B$$

from the first law of thermodynamics, we further have

$$U_B - U_A = \Delta Q + P_A V_A - P_B V_B$$

The process is adiabatic, so  $\Delta Q = 0$ . Thus, enthalpy is conserved:

$$H_B = U_B + P_B V_B = U_A + P_A V_A = H_A$$

Because enthalpy and particle number are constant, we can write:

$$dT = \left( \frac{\partial T}{\partial P} \right)_{H,N} dP$$

This partial derivative is defined as the Joule-Thomson coefficient

$$\mu_{JT} = \left( \frac{\partial T}{\partial P} \right)_{H,N} \quad (13.16)$$

We wish to state this factor in terms of the standard set.

Using the Jacobian trick,

$$\begin{aligned} \left( \frac{\partial T}{\partial P} \right)_{H,N} &= \frac{\partial(T, H)}{\partial(P, H)} = \frac{\partial(T, H)}{\partial(P, T)} \cdot \frac{\partial(P, T)}{\partial(P, H)} \\ &= - \left( \frac{\partial H}{\partial P} \right)_{T,N} \cdot \frac{1}{\left( \frac{\partial H}{\partial T} \right)_{P,N}} \end{aligned}$$

We can use the “divide through” trick with the differential (holding  $N$  constant)

$$dH = T dS + V dP$$

to determine these derivatives

$$\begin{aligned} \left( \frac{\partial H}{\partial P} \right)_T &= T \left( \frac{\partial S}{\partial P} \right)_T + V \left( \frac{\partial P}{\partial P} \right)_T = T \left( \frac{\partial S}{\partial P} \right)_T + V \\ \left( \frac{\partial H}{\partial T} \right)_P &= T \left( \frac{\partial S}{\partial T} \right)_P + V \left( \frac{\partial P}{\partial T} \right)_P = T \left( \frac{\partial S}{\partial T} \right)_P = N c_P \end{aligned}$$

Finally, we wish to restate  $\left(\frac{\partial S}{\partial P}\right)_T$  in terms of the standard set. Using a Maxwell relation, the differential

$$dG = -S dT + V dP + \mu dN$$

shows us

$$-\left(\frac{\partial S}{\partial P}\right)_{T,N} = \left(\frac{\partial V}{\partial T}\right)_{P,N} = V\alpha$$

Thus, combining all of these,

$$\begin{aligned}\mu_{JT} &= -(-TV\alpha + V) \cdot \frac{1}{Nc_P} \\ &= \frac{V}{Nc_P}(T\alpha - 1)\end{aligned}$$

## 13.6 Heat Capacity

We can write  $c_V$  as:

$$\begin{aligned}c_V &= \frac{T}{N} \left(\frac{\partial S}{\partial T}\right)_V = \frac{T}{N} \frac{\partial(S, V)}{\partial(T, N)} \\ &= \frac{T}{N} \frac{\partial(S, V)}{\partial(P, T)} \frac{\partial(P, T)}{\partial(T, V)} \\ &= -\frac{T}{N} \frac{\partial(S, V)}{\partial(P, T)} \frac{1}{\frac{\partial(V, T)}{\partial(P, T)}} \\ &= \frac{T}{N} \left[ \left(\frac{\partial S}{\partial P}\right)_{T,N} \left(\frac{\partial V}{\partial T}\right)_{P,N} - \left(\frac{\partial S}{\partial T}\right)_{P,N} \left(\frac{\partial V}{\partial P}\right)_{T,N} \right] \frac{1}{\left(-\frac{\partial V}{\partial P}\right)_{T,N}} \\ &= \frac{T}{N} \left[ -\left(\frac{\partial V}{\partial T}\right)_{P,N} (V\alpha) - \left(\frac{Nc_P}{T}\right) (-V\kappa_T) \right] \left(\frac{1}{V\kappa_T}\right) \\ &= \frac{T}{N} \left[ (-V\alpha)(V\alpha) - \left(\frac{Nc_P}{T}\right) (-V\kappa_T) \right] \left(\frac{1}{V\kappa_T}\right) \\ &= c_P - \frac{\alpha^2 TV}{N\kappa_T}\end{aligned}$$

The last equation

$$c_P = c_V + \frac{\alpha^2 TV}{N\kappa_T} \tag{13.17}$$

is more general than the Meyer relation, and additionally shows that because it has 4 standard derivatives, only 3 of them are independent.

## 13.7 General Strategy

A general strategy may be outlined as follows:

1. Express partial derivative as Jacobian
2. Insert (usually)  $\partial(P, T)$  using chain rule, and manipulate reciprocals
3. Eliminate  $F, G, H, U$  using known derivative or differential form
4. If extensive, move  $\mu$  to the numerator and eliminate using the Gibbs-Duhem relation
5. Move  $S$  to the numerator and eliminate it with specific heat (if derivative wrt  $T$ ) or maxwell relation (wrt  $P$ )
6. Move  $V$  to the top and eliminate with  $\alpha$  or  $\kappa_T$
7. Eliminate  $c_V$  with  $c_P$ .

## 13.8 Examples

### 13.8.1 Example A

$$\left(\frac{\partial V}{\partial P}\right)_{S,N}$$

Applying Jacobian:

$$\frac{\partial(V, S, N)}{\partial(P, S, N)} = \frac{\partial(V, S, N)}{\partial(P, T, N)} \frac{\partial(P, T, N)}{\partial(P, S, N)} = \frac{\partial(V, S, N)}{\partial(P, T, N)} \frac{\partial(T, P, N)}{\partial(S, P, N)}$$

The second term becomes

$$\left[\left(\frac{\partial S}{\partial T}\right)_{P,N}\right]^{-1} = \frac{T}{Nc_P}$$

Expanding the first term,

$$\begin{aligned} \frac{\partial(V, S, N)}{\partial(P, T, N)} &= \left(\frac{\partial V}{\partial P}\right)_{T,N} \left(\frac{\partial S}{\partial T}\right)_{P,N} - \left(\frac{\partial V}{\partial T}\right)_{P,N} \left(\frac{\partial S}{\partial P}\right)_{T,N} \\ &= (-V\kappa_T)(Nc_P/T) - (V\alpha) \left(\frac{\partial S}{\partial P}\right)_{T,N} \end{aligned}$$

The final term can be solved using a maxwell relation with the differential

$$-S dT + V dP + \mu dN$$

thus,

$$\left(\frac{\partial S}{\partial P}\right)_{T,N} = -\left(\frac{\partial V}{\partial T}\right)_{P,N} = -V\alpha$$

Plugging in, we obtain

$$\left(\frac{\partial V}{\partial P}\right)_{S,N} = -V\kappa_T + \frac{V^2 T \alpha^2}{N c_P}$$

### 13.8.2 Example C

$$\left(\frac{\partial F}{\partial S}\right)_{P,N}$$

Recall  $dF = -S dT - P dV + \mu dN$ . Thus,

$$\begin{aligned} \left(\frac{\partial F}{\partial S}\right)_{P,N} &= -S \left(\frac{\partial T}{\partial S}\right)_{P,N} - P \left(\frac{\partial V}{\partial S}\right)_{P,N} + \mu \left(\frac{\partial N}{\partial S}\right)_{P,N} \\ \left(\frac{\partial F}{\partial S}\right)_{P,N} &= -\frac{T}{N c_P} - P \left(\frac{\partial V}{\partial S}\right)_{P,N} \end{aligned}$$

Using the jacobian trick,

$$\left(\frac{\partial V}{\partial S}\right)_{P,N} = \left(\frac{\partial V}{\partial T}\right)_{P,N} \left(\frac{\partial S}{\partial T}\right)_{P,N}^{-1} = \frac{V T \alpha}{N c_P}$$

Thus,

$$\left(\frac{\partial F}{\partial S}\right)_{P,N} = -\frac{T(1 + P V \alpha)}{N c_P}$$

### 13.8.3 Example E

$$\left(\frac{\partial T}{\partial N}\right)_{S,V}$$

Using a maxwell relation,

$$\begin{aligned} T dS - P dV + \mu dN \\ \left(\frac{\partial T}{\partial N}\right)_{S,V} &= \left(\frac{\partial \mu}{\partial S}\right)_{V,N} \end{aligned}$$

Using the Gibbs-Duhem relation,

$$\begin{aligned} d\mu &= -\left(\frac{S}{N}\right) dT + \left(\frac{V}{N}\right) dP \\ \left(\frac{\partial \mu}{\partial S}\right)_{V,N} &= -\frac{V}{N} \left(\frac{\partial T}{\partial S}\right)_{V,N} + \frac{V}{N} \left(\frac{\partial P}{\partial S}\right)_{V,N} \\ &= -\frac{S}{N} \frac{T}{N c_V} + \frac{V}{N} \left(-\alpha V + \frac{N c_P \kappa_T}{\alpha T}\right)^{-1} \end{aligned}$$



# Chapter 14

## Extremum Principles

We saw earlier that entropy is maximised in thermal equilibrium; this is known as an *extremum principle*.

### 14.0.1 Revisited

If the movable piston is adiabatic (i.e. heat cannot flow through), the system is actually indeterminate! There is no unique equilibrium; it oscillates. However, if a correction term is added for the average velocity of the particles being adjusted for the velocity of the piston, the oscillations become damped and the equilibrium can be found.

## 14.1 Energy Minimum Principle

Consider

$$S(U, X)$$

for a thermally isolated system, where  $X$  is an extensive parameter which describes the distribution of amount of something in a composite system. For example,  $X$  could be  $N$ ,  $U$ ,  $V$  of a subsystem.

From the maximisation of entropy at equilibrium, we then must have

$$\left(\frac{\partial S}{\partial X}\right)_U = 0 \quad \left(\frac{\partial^2 S}{\partial X^2}\right)_U < 0$$

Thus, we can expand the entropy around leading order terms as

$$dS \approx \left(\frac{\partial S}{\partial U}\right)_U dU + \frac{1}{2} \left(\frac{\partial^2 S}{\partial X^2}\right)_U (dX)^2$$

We know that

$$\left(\frac{\partial S}{\partial U}\right)_X = \frac{1}{T}$$

allowing us to write

$$dS \approx \frac{dU}{T} + \frac{1}{2} \left( \frac{\partial^2 S}{\partial X^2} \right)_U (dX)^2 \quad (14.1)$$

Considering a quasi-static process with no heat exchange, the entropy remains constant. Thus, work must be added to the energy to maintain equilibrium.

$$dU = T dS - \frac{T}{2} \left( \frac{\partial^2 S}{\partial X^2} \right)_U (dX)^2 \quad (14.2)$$

from this, we see that because there is no  $dX$  term,

$$\left( \frac{\partial U}{\partial X} \right)_S = 0$$

Additionally, we see that

$$\left( \frac{\partial^2 U}{\partial X^2} \right)_S = -T \left( \frac{\partial^2 S}{\partial X^2} \right)_U > 0$$

thus, energy is minimized at equilibrium for a constant entropy. Consequently, the maximum amount of work for  $dS, dN = 0$  is  $dW = -dU$

### 14.1.1 Revisited

The system once again has no heat exchanged with the environment, but the system can do work (quasistatically) via the movable piston between the two subsystems. We instead examine this system via *implicit functions*. We hold entropy fixed, so

$$S(X, U, Y) = C$$

creating an implicit function  $U(X, Y)$ . The differential

$$0 = dS = \left( \frac{\partial S}{\partial X} \right)_{U,Y} dX + \left( \frac{\partial S}{\partial U} \right)_{X,Y} dU + \left( \frac{\partial S}{\partial Y} \right)_{X,U} dY$$

Holding  $dY = 0$ , we can then divide through by  $dX$ , yielding

$$0 = \left( \frac{\partial S}{\partial X} \right)_{U,Y} + \left( \frac{\partial S}{\partial U} \right)_{X,Y} \left( \frac{\partial U}{\partial X} \right)_{S,Y}$$

yielding

$$\left( \frac{\partial U}{\partial X} \right)_{S,Y} = - \left( \frac{\partial S}{\partial X} \right)_{U,Y} / \left( \frac{\partial S}{\partial U} \right)_{X,Y} \quad (14.3)$$

Because  $S$  is constant wrt  $X$  by constraint of the problem (note however, that  $S$  is not necessarily constant wrt  $U$ ), we see that

$$\left( \frac{\partial U}{\partial X} \right)_{S,Y} = 0 \quad (14.4)$$

To further show that this is a minimum, we define

$$\phi \equiv \left( \frac{\partial U}{\partial X} \right)_{S,Y}$$

Thus,  $\phi = \phi(X, U)$ . Then,

$$\left( \frac{\partial^2 U}{\partial x^2} \right)_{S,Y} = \left( \frac{\partial \phi}{\partial X} \right)_{S,Y} = \left( \frac{\partial \phi}{\partial U} \right)_{X,Y} \left( \frac{\partial U}{\partial X} \right)_{S,Y} + \left( \frac{\partial \phi}{\partial X} \right)_{U,Y} \left( \frac{\partial X}{\partial X} \right)_{S,Y}$$

Plugging in  $\phi = 0$ , we see that

$$\begin{aligned} \left( \frac{\partial^2 U}{\partial X^2} \right)_{S,Y} &= \left( \frac{\partial \phi}{\partial X} \right)_{U,Y} \\ &= \frac{\partial}{\partial x} \left[ - \left( \frac{\partial S}{\partial X} \right)_{U,Y} / \left( \frac{\partial S}{\partial U} \right)_{X,Y} \right]_{U,Y} \\ &= \vdots \\ &= -T \left( \frac{\partial^2 S}{\partial X^2} \right)_{U,Y} > 0 \end{aligned}$$

## 14.2 Minimum Principle for Helmholtz Free Energy

Consider a system in contact with a thermal reservoir (e.g. the atmosphere). Temperature is then constant, while energy and entropy may be exchanged. We may derive this principle in two ways: minimize energy with constant entropy of system+reservoir, or maximise entropy with constant energy of system+reservoir.

Consider the former. The total entropy should not change:

$$\frac{\partial}{\partial X} (S + S_R) = 0 \implies \frac{\partial S}{\partial X} = -\frac{\partial S_R}{\partial X}, \quad \frac{\partial^2 S}{\partial X^2} = -\frac{\partial^2 S_R}{\partial X^2}$$

but the total energy is at a minimum:

$$\frac{\partial}{\partial X} (U + U_R) = 0 \quad \frac{\partial^2}{\partial X^2} (U + U_R) > 0$$

Because only heat can be exchanged between the two subsystems, we further impose

$$dU_R = T_R dS_R \implies \frac{\partial U_R}{\partial X} = T_R \frac{\partial S_R}{\partial X}$$

Plugging into the derivative of Helmholtz free energy, we find that

$$\frac{\partial F}{\partial X} = \frac{\partial}{\partial X} (U - TS) = -\frac{\partial}{\partial X} (U_R - T_R S_R) = 0$$

We can bind a similar result for

$$\frac{\partial^2 F}{\partial X^2} = \frac{\partial^2}{\partial X^2}(U - TS) > -\frac{\partial^2}{\partial X^2}(U_R - T_R S_R) = 0$$

Thus, we have the conditions

$$\begin{aligned}\frac{\partial F}{\partial X} &= 0 \\ \frac{\partial^2 F}{\partial X^2} &> 0\end{aligned}$$

or,  $F$  is at a minimum.

Similarly, we can examine the other setup—maximising entropy at a fixed energy. This proof is very similar to the previous and is left as an exercise.

The maximum amount of work that can be done to the system is given

$$dF = -SdT - PdV + \mu dN = dW$$

thus, the maximum amount of work the system can do is given

$$-dW = -dF \tag{14.6}$$

### 14.3 Minimum Principle for Enthalpy

When pressure is held constant, equilibrium occurs when enthalpy is minimized. The proof is similar to that of Helmholtz free energy, albeit with pressure reservoir. We hold the only work that can be done is through  $PV$  work

$$dU_R = -P_R dV_R \implies \frac{\partial U_R}{\partial X} = -P_R \frac{\partial V_R}{\partial X}$$

and either hold entropy constant and minimize energy, or hold energy constant and maximise entropy. Holding the total volume constant

$$\frac{\partial}{\partial X}(V + V_R) = 0$$

The derivative of enthalpy is given

$$\frac{\partial H}{\partial X} = \frac{\partial}{\partial X}(U + P_R V) = \frac{\partial}{\partial X}(U + U_R) = 0$$

and the second derivative as

$$\frac{\partial^2 H}{\partial X^2} = \frac{\partial^2}{\partial X^2}(U + P_R V) = \frac{\partial^2}{\partial X^2}(U + U_R) > 0$$

Further, we have

$$dH = T dS + V dP + \mu dN = dQ$$

so the maximum amount of heat that can be released is

$$-dQ = -dH \tag{14.7}$$

The enthalpy is often useful in chemistry, as most reactions are done at atmospheric pressure, and may be referred to as the “heat content”

## 14.4 Minimum Principle for Gibbs Free Energy

The Gibbs free energy is also minimized at equilibrium. Again, the proof is similar to the Helmholtz free energy and Enthalpy.

$$dG = -S dT + V dP + \mu dN = -S dT$$

### 14.4.1 Chemical Reactions

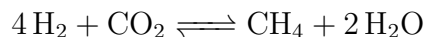
For a system of  $r$  components, the Gibbs free energy is given

$$G = \sum_j^r \mu_j N_j$$

A chemical reaction can be written, in terms of integral stoichiometric coefficients can be written

$$0 \rightleftharpoons \sum_j \nu_j A_j \quad (14.8)$$

for chemical species  $A_j$ . For example, the reaction



has coefficients

$$\nu_{\text{CH}_4} = 1 \quad \nu_{\text{H}_2\text{O}} = 2 \quad \nu_{\text{H}_2} = -4 \quad \nu_{\text{CO}_2} = -1$$

Each of the chemical species can only change in number through stoichiometric reaction, so we have

$$\frac{dN_i}{\nu_i} = d\mathcal{N}$$

Thus, at constant temperature and pressure, the Gibbs free energy becomes

$$dG = dN \sum_j^r \mu_j \nu_j = 0$$

If the system starts out of equilibrium at  $N_j^0$ , the equilibrium number will become

$$N_j = N_j + \nu_j \Delta\mathcal{N}$$

Because the number of particles must be positive, there must be a maximum and minimum  $\Delta\mathcal{N}$ . From this, we define a degree of reaction

$$\varepsilon \equiv \frac{\Delta\mathcal{N} - \Delta\mathcal{N}_{\min}}{\Delta\mathcal{N}_{\max} - \Delta\mathcal{N}_{\min}} \quad (14.9)$$

From chemistry we know that at constant  $T, P$ , the change in Gibbs free energy is given

$$\Delta G = \Delta H - T \Delta S \quad (14.10)$$

The values  $\Delta H$  and  $\Delta S$  can be obtained from a table of standard values. At equilibrium,  $\Delta G = 0$ , and  $\Delta G < 0$  means that the reaction will occur in the forward direction and  $\Delta G > 0$  means the reaction will occur in the reverse direction.

Returning to the previous chemical reaction, we start with 5 mol  $\text{H}_2$ , 1 mol  $\text{H}_2$ , 1 mol  $\text{CH}_4$ , and 3 mol  $\text{H}_2\text{O}$ . The new component numbers are given:

$$N_{\text{H}_2} = 5 - 4\Delta\mathcal{N}$$

$$N_{\text{H}_2\text{O}} = 3 + 2\Delta\mathcal{N}$$

$$N_{\text{CH}_4} = 1 + \Delta\mathcal{N}$$

$$N_{\text{CO}_2} = 1 - \Delta\mathcal{N}$$

The limits are defined by the depletion of  $\text{CO}_2$  with  $\Delta\mathcal{N}_{\text{max}} = 0$  and the depletion of  $\text{CH}_4$  with  $\Delta\mathcal{N}_{\text{min}} = 0$ . Thus, the degree of reaction for this system is

$$\varepsilon = \frac{1}{2}(1 + \Delta\mathcal{N})$$

# Chapter 15

## Stability Conditions

A system is stable if it responds to any small perturbation by trying to return to its original configuration. In this way, we can view it as the minimum of some sort of potential well. We will examine stability with extremum principles; while it may seem like we are examining a very specific case, the results obtained are actually general.

### 15.1 Extensive Stability Conditions

Consider two thermodynamic systems, denoted  $S(U, V, N)$  and  $\hat{S}(\hat{U}, \hat{V}, \hat{N})$ . We will combine these two systems to form a composite system, isolated from the rest of the universe. Similar to the previous chapter, we will perturb an extensive quantity  $X$ .

#### 15.1.1 Volume and Energy

Assume the partition between the system is a moveable wall, such that

$$V_T = V + \hat{V}$$

By the minimum energy principle, we know that

$$\Delta U_T = U(S, V + \Delta V, N) + \hat{U}(\hat{S}, \hat{V} - \Delta V, \hat{N}) - U(S, V, N) - \hat{U}(\hat{S}, \hat{V}, \hat{N}) \geq 0$$

Note that equality holds only when  $\Delta V = 0$ . Assuming the two systems are identical, and the perturbation  $\Delta V \ll 1$

$$U(S, V + \Delta V, N) + U(S, V - \Delta V, N) - 2U(S, V, N) \geq 0$$

Dividing by  $\Delta V^2$  and taking the limit  $\Delta V \rightarrow 0$ , this expression, we see that this is the finite-difference expression for the second order derivative:

$$\left( \frac{\partial^2 U}{\partial V^2} \right)_{S, N} \geq 0 \tag{15.1}$$

or, in terms of the isentropic compressibility (an analogue to the isothermal compressibility)

$$\kappa_S = -\frac{1}{V} \left( \frac{\partial V}{\partial P} \right)_{S,N}$$

$$\left( \frac{\partial^2 U}{\partial V^2} \right)_{S,N} = - \left( \frac{\partial P}{\partial V} \right)_{S,N} = \frac{-1}{\left( \frac{\partial V}{\partial P} \right)_{S,N}}$$

Thus,

$$\frac{1}{V \kappa_S} \geq 0 \quad (15.2)$$

further, because  $V > 0$ , we have  $\kappa_S > 0$ , or an increase in pressure always leads to a decrease in volume.

### 15.1.2 Heat and Energy

Instead of varying volume, we instead vary entropy. Similar to the previous section, we obtain

$$\left( \frac{\partial^2 U}{\partial S^2} \right)_{V,N} \geq 0 \quad (15.3)$$

In terms of  $c_V$ ,

$$\left( \frac{\partial^2 U}{\partial S^2} \right)_{V,N} = \left( \frac{\partial T}{\partial S} \right)_{V,N} = \frac{T}{N c_V}$$

thus,

$$\frac{T}{N c_V} \geq 0 \quad (15.4)$$

Or, when we add heat to a system, the temperature increases.

### 15.1.3 Helmholtz

We have a composite system in contact with a thermal reservoir. Thus, the Helmholtz Free Energy is minimized. Perturbing volume, as above, we obtain

$$\left( \frac{\partial^2 F}{\partial V^2} \right)_{T,N} \geq 0 \quad (15.5)$$

thus,

$$\left( \frac{\partial^2 F}{\partial V^2} \right)_{T,N} = - \left( \frac{\partial P}{\partial V} \right)_{T,N} \geq 0$$

or

$$\frac{1}{V \kappa_T} \geq 0 \quad (15.6)$$

so

$$\kappa_T \geq 0 \quad (15.7)$$



### 15.1.4 Enthalpy

Consider a composite system at a constant volume. Perturbing entropy,

$$\left(\frac{\partial^2 H}{\partial S^2}\right)_{P,N} \geq 0 \quad (15.8)$$

thus,

$$\left(\frac{\partial^2 H}{\partial S^2}\right)_{P,N} = \left(\frac{\partial T}{\partial S}\right)_{P,N} = \frac{T}{Nc_P}$$

or

$$\frac{T}{Nc_P} \geq 0 \quad (15.9)$$

so

$$c_P \geq 0 \quad (15.10)$$

### 15.1.5 Inequalities

Recall that

$$c_P = c_V + \frac{\alpha^2 TV}{N\kappa_T}$$

from the stability conditions, we see that all terms are positive. Thus,

$$c_P \geq c_V \quad (15.11)$$

Similarly, the relation

$$\kappa_S = \kappa_T - \frac{\alpha^2 TV}{Nc_P} \quad (15.12)$$

shows us that

$$\kappa_T \geq \kappa_S \quad (15.13)$$

## 15.2 Intensive Stability Conditions

Take the temperature of Hemholtz free energy. We have

$$\left(\frac{\partial F}{\partial T}\right)_{V,N} = -S$$

so

$$\left(\frac{\partial^2 F}{\partial T^2}\right)_{V,N} = -\left(\frac{\partial S}{\partial T}\right)_{V,N} = -\frac{Nc_V}{T} \leq 0$$

in general, the extensive equilibrium conditions are positive, while the intensive equilibrium conditions are negative. We can alternatively derive this using heat transfer at minimum energy:

$$\left(\frac{\partial U}{\partial S}\right)_{V,N} = T \implies \left(\frac{\partial^2 U}{\partial S^2}\right)_{V,N} = \left(\frac{\partial T}{\partial S}\right)_{V,N} \geq 0$$

Thus,

$$\left(\frac{\partial^2 F}{\partial T^2}\right)_{V,N} = -\left(\frac{\partial T}{\partial S}\right)_{V,N}^{-1} = -\left(\frac{\partial^2 U}{\partial S^2}\right)_{V,N}^{-1} \geq 0$$

Notice that this is a relation between the conjugate variable  $T, S$  between the legendre transforms  $F, U$ . The legendre transform thus gives us a way to convert the stability conditions for extensive quantities to those for intensive quantities.

# Chapter 16

## Phase Transitions

A phase transition is an abrupt change in a certain property of a material; “analyticity” is no longer applicable at the transition. There are two types of transitions. *First order* transitions are characterized by a discontinuous first derivative of free energy and a discontinuous change in a variable  $X$ . A *Second Order* transition has this discontinuity at a higher order derivative, and consequently has a continuous change in the variable  $X$ .

### 16.1 van der Waals Fluid

A classical ideal gas has no phase transition; this is due to the lack of interactions between particles. By adding a potential, we can describe phase transitions. Typical potentials have repulsions at short ranges (Pauli repulsion force from overlapping orbitals) and attraction at larger distances (dipole-dipole interactions). We begin by examining the Helmholtz Free Energy of a classical ideal gas:

$$F_{IG}(T, V, N) = -Nk_B T \left[ \ln \left( \frac{V}{N} \right) + \frac{3}{2} \ln(k_B T) + X \right]$$

we can add additional terms to encapsulate the interactions. The effective volume is reduced due to the repulsion by a effective factor of  $b$  per particle, while an attraction energy that reduces the total energy is proportional to the density of particles  $N/V$  and the number of particles  $N$  is approximated by  $aN^2/V$ . Thus, the Free Energy becomes

$$F_{vdW} = -Nk_B T \left[ \ln \left( \frac{V - bN}{N} \right) + \frac{3}{2} \ln(k_B T) + X \right] - a \left( \frac{N^2}{V} \right) \quad (16.1)$$

for positive constants  $a, b > 0$ .

Computing the derivatives of the Free energy wrt  $T, V$ , we obtain the equations of state

$$P = \frac{Nk_B T}{V - bN} - \frac{aN^2}{V^2} \quad (16.2a)$$

$$U = \frac{3}{2}k_B T - \frac{a}{N^2}V \quad (16.2b)$$

Energy from vdW

The derivative gives:

$$\begin{aligned}
 S &= -\frac{1}{T} \left( F + \frac{aN^2}{V} \right) + \frac{3}{2}k_B \\
 &= -\frac{1}{T}(U - TS) - \frac{aN^2V}{T} + \frac{3}{2}Nk_B \\
 &= S - \frac{U}{T} - \frac{aN^2}{VT} + \frac{3}{2}Nk_B U = \frac{3}{2}Nk_B T - \frac{aN^2}{V}
 \end{aligned}$$

Because surfaces and other interfaces are neglected, the vdW fluid is an extensive system.

At high temperatures, the vdW fluid approaches the behaviour of an ideal gas. However, it deviates from an ideal gas when the temperature decreases. At high temperature, a  $PV$  plot is monotonic decreasing, but at a critical temperature  $T_c$ , the  $PV$  plot will have an inflection point

$$\left( \frac{\partial P}{\partial V} \right)_{T,N} = 0 \quad \left( \frac{\partial^2 P}{\partial V^2} \right)_{T,N}$$

This inflection point is characterized with the following:

$$V_c = 3bN \quad P_c = \frac{a}{27b^2} \quad k_B T_c = \frac{8a}{27b} \quad (16.3)$$

These three values can be determined from two equations because only two are independent. They satisfy the relation

$$\frac{P_c V_c}{N k_B T_c} = \frac{3}{8} \quad (16.4)$$

Normalizing the variables with  $\tilde{V} = V/V_c$  and so forth, the vdW equation can be rewritten:

$$\left( \tilde{P} + 3\tilde{V}^{-2} \right) \left( 3\tilde{V} - 1 \right) = 8\tilde{T} \quad (16.5)$$

### 16.1.1 Instability

If we plot  $PV$  for  $T < T_c$ , we see that for a certain region, the volume increases in response to an increase in temperature; however, this violates the stability condition  $\kappa_T \geq 0$  we showed in the previous chapter. This section of the curve represents unphysical states.

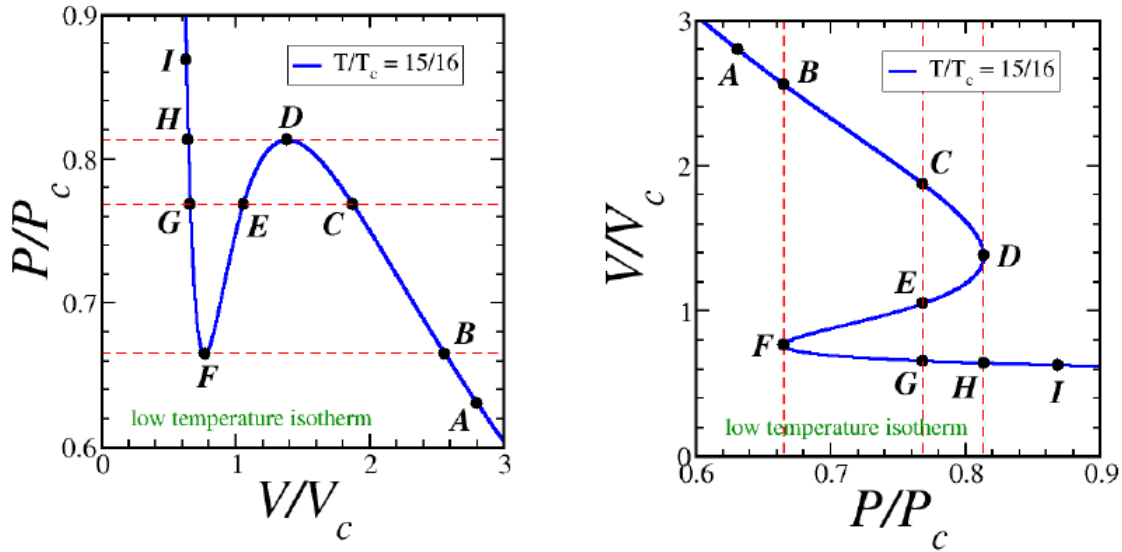


Figure 16.1:  $PV$

Instead, we plot  $VP$ , we get a triple-valued function in the region where  $\frac{\partial V}{\partial P} \geq 0$ .

Rewriting one of the equations of state for the van der Waals fluid,

$$(PV^2 + aN^2)(V - bN) = Nk_BTV^2 \quad (16.6)$$

yields a cubic equation with three solutions. One of these solutions is the region of the curve where  $\frac{\partial V}{\partial P} \geq 0$ ; this we can discard as we know that it corresponds to an unstable state. Instead we seek to replace this section with a different branch of the cubic.

As we want to vary pressure while holding temperature constant, we choose Gibbs free energy as a potential to study the system:

$$G = \mu N$$

As  $N$  remains constant, we seek to minimise  $\mu$ . As the vdW fluid is extensive, we can use the Gibbs-Duhem relation

$$d\mu = - \left( \frac{S}{N} \right) dT + \frac{V}{N} dP$$

Combining these two expressions,

$$G = \mu N = N \int \frac{V}{N} dP = \int V dP$$

Integrating numerically, we obtain the following graph of  $G_r = [G(P) - G(\frac{1}{3}P_c)]/(V_c P_c)$  against  $\tilde{P}$ .

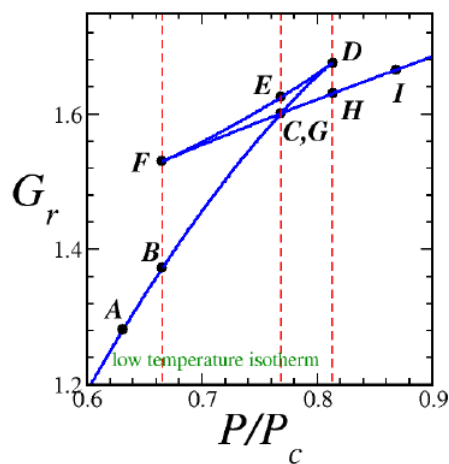


Figure 16.2: GP

Notice that the graph has a crossing point, at  $P \approx 0.76851$ , and a closed loop. The labelled points on the  $G_r \tilde{P}$  graph correspond to the same points on the  $\tilde{V} \tilde{P}$  graph, and show how  $G_r$  is built as  $V$  is integrated. The line CG is the pressure at which the graph crosses, and signifies the transition pressure for this system.

Recall that equilibria occur when  $G$  is minimized; because the loop lies above other portions of the graph, it is not at equilibrium. Thus, the “real” curve  $G$  is the curve in blue shown below

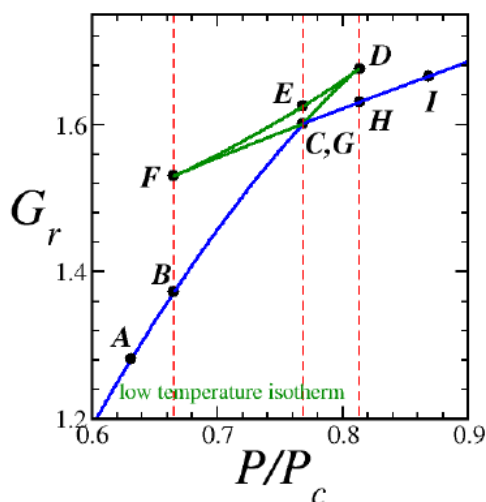


Figure 16.3: Stable

We already noted that the segment  $DF$  is unstable because it violates the stability condition. However, the sections  $CD$  and  $FG$  are long-lived *metastable* states, which are stable to small fluctuations, but large fluctuations cause transitions.

### 16.1.2 Maxwell Construction

While the Gibbs free energy may be obtained, the numerical calculation can be difficult to perform. Instead, the transition temperature may be obtained via the *Maxwell construction*. Extending the line  $CG$  on the  $PV$  curve, the region bounded by  $CDE$  and by  $EFG$  actually have the same area!

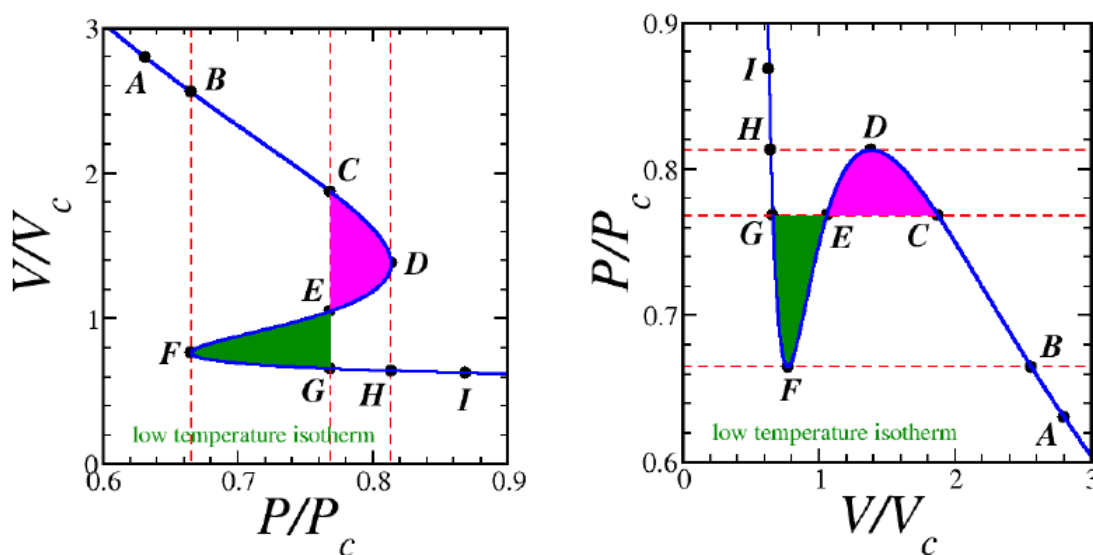


Figure 16.4: Maxwell Construction

This is because this integration follows a closed loop in  $GP$ ; because the starting  $G$  must equal the final, the (signed) sum of the two areas must be zero. This area is conserved if we flip the axes, so we can instead compute  $PdV$ . Guessing  $P$ , we can solve the cubic and obtain endpoints of integration, and adjust our guess according to the difference in the areas.

### 16.1.3 Multiple Phases

Note that points  $C$  and  $G$  exist at the same pressure and at the same Free Energy, but have different volumes; there exists two stable configurations of the fluid at this pressure. Point  $G$ , having a lower volume, corresponds to a liquid phase, while point  $C$  with the greater volume corresponds to a gas. Repeating the maxwell construction for different isotherms, we can determine the transition pressure as a function of temperature.

Plotting the temperature against the volumes of the two phases at the transition, we obtain the below graph. The point at the top of the curve is the *critical point*, characterized by  $T_c$ ,  $V_c$ ,  $P_c$ , and the meta-stable states are so-called “spinodals,” characterized by points  $D$  and  $F$ .

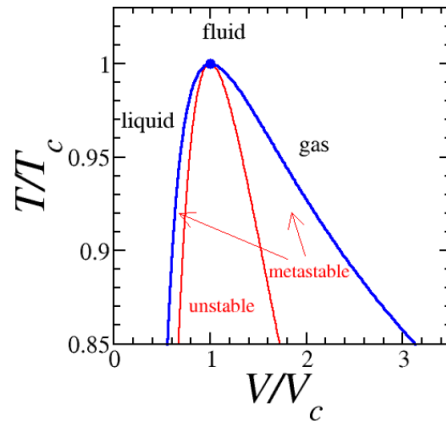


Figure 16.5: idk

If we instead plot the pressure against the temperature, we get obtain the familiar phase diagram. Above  $T_c$ , there is no distinction between liquid and gas; they form a uniform “fluid.” In fact, you can go smoothly from gas to liquid by going around the critical point. The blue line in the phase diagram is known as the *coexistence curve*.

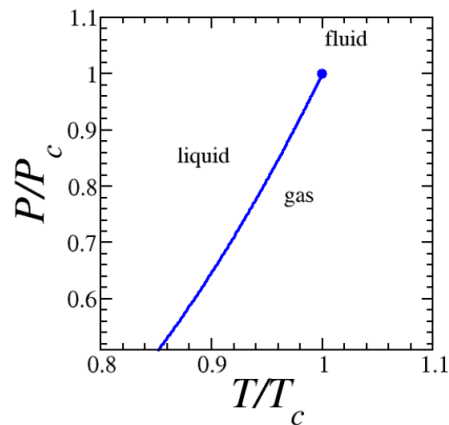


Figure 16.6: Phase Diagram



## 16.2 Helmholtz Free energy

Rewriting the helmholtz free energy in terms of dimensionless quantities,

$$F_r = \frac{F_{vdW}}{V_c P_c} = -\frac{8}{3} \tilde{T} \left[ \ln(3\tilde{V} - 1) + \frac{3}{2} \ln(\tilde{T}) + \tilde{X} \right] - \frac{3}{\tilde{V}} \quad (16.7)$$

The second derivative is then

$$\left( \frac{\partial^2 F_r}{\partial \tilde{V}^2} \right)_{\tilde{T}, N} = \frac{24\tilde{T}}{(3\tilde{V} - 1)^2} - \frac{6}{\tilde{V}^3} \quad (16.8)$$

Because this quantity is related to the isothermal compressibility, it must be positive. If we plot this second derivative against volume for an isotherm, we will see that below the critical temperaure, there is a region where graph goes below zero. The intercepts are the volumes of the liquid and gas respectively, and are connected by a straight  $F'' = 0$  line.

{diagram}

## 16.3 Latent Heat

Energy is required to cross the coexistence curve to go from a liquid to a gas. This energy is known as the *latent heat of vaporisation*. We can obtain the latent heat by examining the entropy. Along an isotherm, the change in entropy is given

$$dS = \left( \frac{\partial S}{\partial V} \right)_{T, V} dV$$

Using the maxwell relation

$$\left( \frac{\partial S}{\partial V} \right)_{T, N} = \left( \frac{\partial P}{\partial T} \right)_{V, N} = \frac{Nk_B}{V - bN}$$

we obtain

$$\Delta S = Nk_B \int_{V_{\text{liquid}}}^{V_{\text{gas}}} \frac{dV}{V - bN}$$

thus, along the isotherm, the change in heat is

$$\Delta Q = T \Delta S$$

The latent heat per particle is defined to be the *latent heat* of the transition

$$\ell = \frac{T \Delta S}{N} \quad (16.9)$$

and the total energy to complete the transition is the *heat capacity* of the transition.

$$L = N\ell = T \Delta S \quad (16.10)$$

## 16.4 Clausius-Clapeyron Equation

The Clausius-Clapeyron Equation relates the slope of the coexistence curve with the volume change and latent heat of the transition.

Let  $X, Y$  denote the liquid phase at two points along the coexistence curve, and  $X', Y'$  denote the gas phase at the corresponding points along the coexistence curve. The pairs  $X, X'$  and  $Y, Y'$  are very close, so  $dT = dT'$  and  $dP = dP'$ .

Further, because the liquid and gas phase are at equilibrium,  $\mu_X = \mu_{X'}$  and similarly at  $Y$ . Using Gibbs-Duhem and the same  $N$  number of particles for both transitions,

$$d\mu = -\frac{S}{N} dT + \frac{V}{N} dP$$

$$d\mu' = -\frac{S'}{N} dT + \frac{V'}{N} dP$$

Because these two equations are equal, we can obtain:

$$\begin{aligned} \frac{dP}{dT} &= \frac{S' - S}{V' - V} = \frac{\Delta S}{\Delta V} = \frac{T\Delta S}{T\Delta V} \\ \frac{dP}{dT} &= \frac{L}{T\Delta V} \end{aligned} \tag{16.11}$$

where  $\Delta V$  is the volume change of the transition. The Clausius-Clapeyron equation is a general equation for any phase transition. Note that  $L > 0$  and, for most materials,  $\Delta V > 0$ . A notable exception is water, which has  $\Delta V < 0$ , which is why ice floats.

## 16.5 Gibb's Phase Rule

Gibb's phase rule limits the structures allowed in phase diagrams. It specifies the number of independent variables  $F$  of a coexistence structure— $F = 0$  provides a point,  $F = 1$  provides a line,  $F = 2$  provides a surface, and so forth. The rule is given by

$$F = K - \phi + 2 \tag{16.12}$$

where  $K$  is the number of components (e.g. chemical species) and  $\phi$  is the number of coexisting phases.

Let  $x_k^{(j)}$  denote the concentrations (mole fraction) of component  $k$  in phase  $j$ :

$$\{x_k^{(j)} = N_k^{(j)} / N^{(j)} | k = 1, \dots, K\}$$

where

$$N^{(j)} = \sum_k N_k^{(j)}$$

is the total number of all species in phase  $j$ . The concentrations of any given phase must be one:

$$\sum_k x_k^{(j)} = 1 \quad (16.13)$$

which follows immediatly on the definition of  $x_k^{(j)}$ . The number of independent  $x_k^{(j)}$  is then  $K - 1$  which gives a total number of variables of  $P$ ,  $V$  and  $\phi \cdot (K - 1)$ , giving  $2 + \phi \cdot (K - 1)$  variables, which may or may not be independent.

Because the  $\phi$  phases are all in equilibrium (by virtue of coexisting), we further have the  $\phi - 1$  constraints:

$$\mu_k^{(1)} = \cdots = \mu_k^{(\phi)}$$

for each component  $K$ , yielding  $K(\phi - 1)$  total constraints. Subtracting from the total number of variables,

$$F = 2 + \phi \cdot (K - 1) - K(\phi - 1) = 2 + K - \phi$$

For example, for water, at the ice-water transition, we have

$$F = 2 + (K = 1) - (\phi = 2) = 1$$

as there is only 1 component ( $\text{H}_2\text{O}$ ), but two coexistant phases.



# Chapter 17

## Nernst Postulate

Recall that the Nernst Postulate supposes that the entropy of a thermodynamic system goes to a constant as the temperature approaches absolute zero; beginning from quantum mechanics, this postulate arises naturally, but in classical mechanics requires it to be true. Most interestingly, it shows the impact of quantum mechanics on the values of macroscopic properties.

### 17.1 Classical Ideal Gas Violation

Recall the entropy of the classical ideal gas. Re-expressing using the equipartition theorem,

$$S(T, V, N) = k_B N \left[ \frac{3}{2} \left( \frac{3}{2} k_B T \right) + \ln \left( \frac{V}{N} \right) + x \right] \quad (17.1)$$

Note that because  $T$  occurs within a natural logarithm, as  $T \rightarrow 0$ ,  $S \rightarrow -\infty$ ; in fact, this occurs for all classical systems. As such, quantum mechanics must be used in the study of thermodynamics

#### 17.1.1 Planck Extension

Planck's extension states that the constant that entropy approaches is zero. While this is often true, there are some exceptions; however, Nernst's Postulate is *always* true.

### 17.2 Consequences

The Nernst postulate imposes constraints on macroscopic quantities at low temperature; namely, specific heat and thermal expansion.

### 17.2.1 Specific Heat

Recall the definition of specific heat as

$$c_X(T) = \frac{T}{N} \left( \frac{\partial S}{\partial T} \right)_{X,N}$$

for  $X = V, P$ . The change in entropy between two temperatures is given

$$S(T_2) - S(T_1) = \int_{T_1}^{T_2} dS = \int_{T_1}^{T_2} \left( \frac{\partial S}{\partial T} \right)_{X,N} dT = \int_{T_1}^{T_2} \frac{Nc_X(T)}{T} dT$$

Suppose, FSOc that  $\lim_{T \rightarrow 0} c_X(T) = c_0 \neq 0$ . Then, for small temperatures,

$$\Delta S \approx \int_{T_1}^{T_2} \frac{Nc_0}{T} dT = Nc_0 \ln(T_2/T_1)$$

Fixing  $T_2$  and sending  $T_1 \rightarrow 0$ , we see that  $\Delta S \rightarrow +\infty$ , violating the Nernst Postulate, and leading to a contradiction. Thus, we must have  $c_0 = 0$ , and specific heats become zero at absolute zero.

Recall for the Classical Ideal Gas,  $c_V = \frac{3}{2}k_B$  is a constant, independent of temperature; this contradicts the previously derived result, and is another example of how the Classical Ideal Gas violates the Nernst Postulate.

### 17.2.2 Thermal Expansion

A similar result can be obtained for the coefficient of thermal expansion:

$$\alpha = \frac{1}{V} \left( \frac{\partial V}{\partial T} \right)_{P,N}$$

We can rewrite this using a maxwell relation on the Gibbs Free Energy, so

$$-V\alpha = \left( \frac{\partial S}{\partial P} \right)_{T,N}$$

Because  $S$  is constant at  $T = 0$ , then this derivative must necessarily be zero at  $T = 0$ . Thus, we must have  $\alpha \rightarrow 0$  as  $T \rightarrow 0$ .

## Part III

# Classical Statistical Mechanics





# Chapter 18

## Statistical Ensembles

A thermodynamic ensemble is a virtual assembly of copies of a thermodynamic system, each representing an accessible microstate of the thermodynamic system, subject to external requirements on the original system; if there are  $g$  accessible microstates, there are  $g$  systems in the ensemble. The three most common ensembles are the microcanonical ensemble (isolated system), canonical (variable energy), and grand canonical ensemble (variable particles and energy).

### 18.1 Microcanonical Ensemble

The Hamiltonian for an isolated system of  $N$  particles with pairwise interactions is

$$H(q, p) = \sum_{j=1}^N \frac{p_j^2}{2m} + \sum_{j=1}^N \sum_{i>j}^N \phi(r_i, r_j) \quad (18.1)$$

To fix the energy, we introduce the constraint

$$\delta(E - H) \quad (18.2)$$

The constant energy defines the *microcanonical ensemble*. Fixing temperature, volume, and particle number, recall that the entropy can be written

$$S(E, V, N) = k_B \ln \Omega(E, V, N) = k_B \ln \left[ \frac{1}{h^{3N} N!} \int dq \int dp \delta(E - H(p, q)) \right] \quad (18.3)$$

This is a  $6N$  dimensional integral which is often difficult, if not impossible to evaluate for interacting systems.<sup>1</sup> Such systems can be approximated using molecular dynamics simulations, where the differential equations

$$\frac{dq_j}{dt} = \frac{p_j}{m}$$

---

<sup>1</sup>The quantity  $\Omega$  is the *microcanonical partition function*. It is sometimes referred to as the *density of states*. Its identity as the microcanonical partition function is why the logarithm yields the potential  $S$ .

$$\frac{dp_j}{dt} = \sum_k F_{jk}$$

lead to evolution

$$q_j(t + \delta t) = q_j(t) + \frac{p_j(t)}{m} \delta t$$

$$p_j(t + \delta t) = p_j(t) + \sum_k F_{jk}(t) \delta t$$

if  $\delta t$  is too large, there is noise due to discretization, and for large  $N$ , there are too many particles to hold in memory.

## 18.2 Canonical Ensemble

The canonical ensemble is held at a constant temperature by contact with a heat reservoir; energy can flow between the ensemble and reservoir through the interface. Thus, the canonical ensemble involves a probability distribution for the energy.

### 18.2.1 Canonical Distribution of Energy

Assume that the system in contact with a thermal reservoir at temperature  $T$ . Because we fix the volume and particle number, the entropy of the reservoir can be written

$$S_R = k_B \ln \Omega_R(E_R)$$

Assuming that the energy of the system+reservoir is isolated from the rest of the universe, we may write

$$E_T = E + E_R$$

From a previous chapter, we determined the probability distribution to be in the form

$$P(E) = \frac{N_T!}{N!N_R!} \frac{\iiint \delta(E - H) \delta(E_R - H_T) dq dp dq_R dp_R}{\iiint \delta(E_T - H_T) dq dp dq_R dp_R}$$

for compactness, we defined a function

$$\Omega_\alpha(E_\alpha) \equiv \frac{1}{h^{3N_\alpha} N_\alpha!} \iint \delta(E_\alpha - H_\alpha) dq_\alpha dp_\alpha$$

thus,

$$P(E) = \frac{\Omega(E) \Omega_R(E_T - E)}{\Omega_T(E_T)}$$

Taking the logarithm,

$$\ln P(E) = \ln \Omega(E) + \ln \Omega_R(E_T - E) - \ln \Omega_T(E_T)$$

With the assumption that for a reservoir,  $E_T > E_R \gg E$ , we Taylor expand the second term to obtain

$$\ln P(E) = \ln \Omega(E) + \ln \Omega_R(E_T) - E \left( \frac{\partial \ln \Omega_R(E_T)}{\partial E} \right) - \ln \Omega_T(E_T) + O[(E/E_T)^2]$$

Because we know the system and reservoir are in thermal equilibrium, we have

$$\frac{\partial \ln \Omega_R(E_T)}{\partial E} = \beta_R = \beta = \frac{1}{k_B T}$$

Defining a quantity  $\ln Z \equiv \ln \Omega_T(E_T) - \ln \Omega_R(E_T)$ , we have

$$\ln P(E) = \ln \Omega(E) - \beta E + \ln(Z)$$

or

$$P(E) = \frac{\Omega(E)}{Z} e^{-\beta E} \quad (18.4)$$

$Z$  is independent of  $E$ , but depends on  $T, V, N$ . We can integrate an expression for  $Z$  by normalizing the probability.

$$Z(T, V, N) = \int \Omega(E, V, N) e^{-\beta E} dE \quad (18.5)$$

$Z$  is the *partition function*<sup>2</sup>. From Equation 18.5, we see that the partition function has the form of the Laplace transform of the function  $\Omega(E)$ .

### 18.2.2 Canonical Distribution in Phase Space

Rather than using the microcanonical distribution of the previous section, we start from phase space considerations. Making the usual assumption of a universal probability distribution over total phase space, the Joint probability distribution for a given energy  $E_T$  is

$$P(q, p, q_R, p_R) = Y \delta(E_T - H - H_R)$$

where  $Y$  is the normalization constant that satisfies

$$\iiint P dq dp dq_R dp_T = 1$$

Thus,

$$Y = \left[ \iiint \delta(E_T - H - H_R) dq dp dq_R dp_T \right]^{-1} = \frac{1}{h^{3N_T} N_T! \Omega_T(E_T)}$$

Integrating over  $\{q_R, p_R\}$  to obtain the marginal distribution,

$$P(q, p) = \frac{h^{3N_R} N_R! \Omega_R(E - H)}{h^{3N_T} N_T! \Omega_T(E_T)} \quad (18.6)$$

---

<sup>2</sup>The  $Z$  stands for the German *Zustandssumme*, or “sum over states”

taking the logarithm and making use of  $E_T \gg H(q, p)$ ,

$$\ln P(q, p) = \ln \Omega_R(E_R) - H(q, p) \left( \frac{\partial \ln \Omega_R(E_T)}{\partial E_T} \right) - \ln \Omega_T(E_T) + \ln \left( \frac{h^{3N_R} N_R!}{h^{3N_T} N_T!} \right)$$

Using a similar definition to  $Z$ , we define another function  $\tilde{Z}$  which yields

$$\ln P(q, p) = \ln(\tilde{Z}) - \beta H(q, p)$$

or

$$P(q, p) = \frac{1}{\tilde{Z}} e^{-\beta H}$$

where normalization yields

$$\tilde{Z} = \iint e^{-\beta H} dp dq$$

Comparing the two expressions, we see

$$\begin{aligned} Z(T, V, N) &= \frac{1}{h^{3N} N!} \iiint \delta(E - H) e^{-\beta E} dE dq dp \\ &= \frac{1}{h^{3N} N!} \iint e^{-\beta H} dq dp \\ &= \frac{1}{h^{3N} N!} \tilde{Z}(T, V, N) \end{aligned}$$

Thus, we define the partition function to be

$$Z(T, V, N) = \frac{1}{h^{3N} N!} \iint e^{-\beta H} dq dp \quad (18.7)$$

We will see that we can obtain all thermodynamic information from the partition function.

### 18.2.3 Liouville Theorem

Recall that from Hamiltonian Mechanics, that any microscopic state of  $N$  particles can be viewed as a point in  $6N$ -dimensional *phase space*, with coordinates of position  $q$  and momentum  $p$ . From the Hamiltonian, the time evolution is given

$$\dot{q}_j = \frac{\partial H}{\partial p_j} \quad \dot{p}_j = -\frac{\partial H}{\partial q_j} \quad (18.8)$$

These microscopic states thus move through phase space along a trajectory defined by these canonical equations. If we consider the density of the points  $\varrho(q, p, t)$ , we are interested in how the density evolves with time.

$$\frac{d\varrho}{dt} = \frac{\partial \varrho}{\partial t} + \sum_{j=1}^{3N} \left[ \frac{\partial \varrho}{\partial q_j} \dot{q}_j + \frac{\partial \varrho}{\partial p_j} \dot{p}_j \right]$$

Because  $\varrho$  is a density of trajectory, it must satisfy the continuity equation, as trajectories can neither begin nor end, and can neither join nor split off into different trajectories:

$$\frac{\partial \rho}{\partial t} + \sum_{j=1}^{3N} \left[ \frac{\partial}{\partial q_j}(\varrho \dot{p}_j) + \frac{\partial}{\partial p_j}(\varrho \dot{q}_j) \right] = 0$$

Expanding the continuity equation,

$$\frac{\partial \varrho}{\partial t} + \varrho \sum_{j=1}^{3N} \left[ \frac{\partial \dot{q}_j}{\partial q_j} + \frac{\partial \dot{p}_j}{\partial p_j} \right] + \sum_{j=1}^{3N} \left[ \frac{\partial \varrho}{\partial q_j} \dot{q}_j + \frac{\partial \varrho}{\partial p_j} \dot{p}_j \right] = 0$$

and substituting the time derivative,

$$\frac{d\varrho}{dt} + \varrho \sum_{j=1}^{3N} \left[ \frac{\partial \dot{q}_j}{\partial q_j} + \frac{\partial \dot{p}_j}{\partial p_j} \right] = 0$$

From hamilton's equations, we have

$$\frac{\partial}{\partial q_j} \dot{q}_j = \frac{\partial}{\partial q_j} \frac{\partial H}{\partial p_j} = \frac{\partial}{\partial p_j} \frac{\partial H}{\partial q_j} = -\frac{\partial \dot{p}_j}{\partial p_j}$$

thus, we obtain Liouville's Theorem,

$$\frac{d\varrho}{dt} = 0 \tag{18.9}$$

Taking the density as the canonical probability distribution  $\varrho = P(H(q, p))$ ,

$$\begin{aligned} 0 &= \frac{d\varrho}{dt} \\ &= \frac{\partial \varrho}{\partial t} + \sum_{j=1}^{3N} \left[ \dot{q}_j \frac{\partial \varrho}{\partial H} \frac{\partial H}{\partial q_j} + \dot{p}_j \frac{\partial \varrho}{\partial H} \frac{\partial H}{\partial p_j} \right] \\ &= \frac{\partial \varrho}{\partial t} + \frac{\partial \varrho}{\partial H} \sum_{j=1}^{3N} [-\dot{q}_j \dot{p}_j + \dot{q}_j \dot{p}_j] \\ &= \frac{\partial \varrho}{\partial t} \end{aligned}$$

thus, the distribution does not change in time. The microcanonical and canonical ensembles depend only on energy and so the probability distributions don't change with time—this leads to stable equilibria.

### 18.2.4 Canonical Distribution Peak

Previously we showed that

$$P(E) = \frac{\Omega(E)}{Z} \exp[-\beta E]$$

The function  $\Omega(E)$  has the form

$$\Omega = \frac{(2\pi m)^{3N/2}}{\Gamma(3N/2)} E^{3N/2-1}$$

For most systems, we see that

$$\Omega \propto E^f$$

for a function  $f$  that varies slowly for  $E$  and can be approximated as a constant. Finding the peak by

$$0 = \frac{\partial}{\partial E} \ln P = \frac{f}{E} - \beta$$

we see the equilibrium peak is

$$E_{eq} = f k_B T$$

The width may be approximated to a gaussian as

$$-\frac{1}{\sigma_E^2} = \left. \frac{\partial^2 \ln P}{\partial E^2} \right|_{E=E_{eq}} = -\frac{f}{E_{eq}^2}$$

or, the relative width is given

$$\frac{\sigma_E}{E} = \frac{1}{\sqrt{f}}$$

Generally,  $f \propto N$ , so for a mole, the relative width is on the order of  $10^{-11}$ .

## 18.3 Partition Function and Free Energy

The canonical probability distribution is given

$$\ln P = -\beta E + \ln \Omega - \ln Z$$

Using the fact that  $S = k_B \ln \Omega$ , we see that

$$\begin{aligned} \ln Z &= -\beta E + S/k_B - \ln P \\ &= -\beta E + TS/k_B T - \ln P \\ &= -\beta(E - TS) - \ln P \\ &= -\beta F - \ln P \end{aligned}$$

However, the peak height of  $P$  is proportional to a power of  $N$ , so  $\ln P \sim O(\ln N)$ . As an extensive parameter, the Helmholtz free energy is  $F \sim O(N)$ , so for  $N \sim 10^{20}$ ,

$$\frac{\ln N}{N} \approx 5 \times 10^{-19}$$

which is negligible. Thus,

$$F = -k_B T \ln Z \tag{18.10}$$

Additionally, using  $\Omega = e^{S/k_B}$

$$\Omega \cdot e^{-\beta E} = e^{-\beta F} \tag{18.11}$$

### 18.3.1 Identities

There are numerous thermodynamic identities which can be derived from statistical mechanics.

$$\begin{aligned}
 \frac{\partial}{\partial \beta} \ln Z &= \frac{1}{Z} \frac{\partial Z}{\partial \beta} \\
 &= \frac{1}{Z} \frac{\partial}{\partial \beta} \left[ \frac{1}{h^{3N} N!} \int dq \int dp \exp[-\beta H] \right] \\
 &= -\frac{1}{Z} \cdot \frac{1}{h^{3N} N!} \int dq \int dp H \exp[-\beta H] \\
 &= -\frac{\int dq \int dp H \exp[-\beta H]}{\int dq \int dp \exp[-\beta H]} \\
 &= \langle E \rangle = U
 \end{aligned}$$

or,

$$\frac{\partial}{\partial \beta}(\beta F) = U \quad (18.12)$$

As statistical mechanics takes into account microscopic effects, it can be used to show relationships inaccessible to thermodynamics. Using the fact that

$$\frac{d\beta}{dT} = -\frac{1}{k_B T^2}$$

we return to the definition

$$c_V = \frac{T}{N} \left( \frac{\partial S}{\partial T} \right)_{V,N}$$

At fixed  $V, N$ , we divide through  $dU + T dS$  to obtain

$$c_V = \frac{1}{N} \left( \frac{\partial U}{\partial T} \right)_{V,N} = \frac{1}{N} \left( \frac{\partial U}{\partial \beta} \right)_{V,N} \frac{d\beta}{dT} = -\frac{1}{N k_B T^2} \left( \frac{\partial \langle E \rangle}{\partial \beta} \right)_{V,N}$$

Taking the derivative

$$\begin{aligned}
 \frac{\partial \langle E \rangle}{\partial \beta} &= \frac{\partial}{\partial \beta} \left[ \frac{\int dq \int dp H \exp[-\beta H]}{\int dq \int dp \exp[-\beta H]} \right] \\
 &= -\frac{\int dq \int dp H^2 \exp[-\beta H]}{\int dq \int dp \exp[-\beta H]} + \left[ \frac{\int dq \int dp H \exp[-\beta H]}{\int dq \int dp \exp[-\beta H]} \right]^2 \\
 &= -\langle E^2 \rangle + \langle E \rangle^2
 \end{aligned}$$

yielding

$$c_V = \frac{1}{N k_B T^2} \left( \langle E^2 \rangle - \langle E \rangle^2 \right) \quad (18.13)$$

which supposedly cannot be obtained through thermodynamics alone. This above equation is the most accurate way to compute the specific heat through computational methods. Similar expressions exist for other quantities.

## 18.4 Momentum-Independent Interactions

Computation of the partition function simplifies when forces do not depend on momenta, which applies to large systems, but notably not particles in magnetic fields. For momentum-independent interactions, we can easily integrate over momentum to obtain:

$$\begin{aligned} Z &= \frac{1}{h^{3N} N!} \int dq \int dp \exp \left[ -\beta \left( \sum_j \frac{p_j^2}{2m} + \sum_j \sum_{i>j} \phi(r_i, r_j) \right) \right] \\ &= \frac{(2\pi m k_B T)^{3N/2}}{h^{3N} N!} \int dq \exp \left[ -\beta \sum_j \sum_{i>j} \phi(r_i, r_j) \right] \end{aligned}$$

For the classical ideal gas, there are no interparticle interactions, so  $\phi = 0$ , so

$$Z_{ideal} = \frac{(2\pi m k_B T)^{3N/2} V^N}{h^{3N} N!} \quad (18.14)$$

Computing the Helmholtz free energy from this partition function agrees to the previously determined expression to within an additive constant.

### 18.4.1 Markov-chain Monte Carlo Simulations

In general, the integral of the interactions are typically obtained numerically. Monte Carlo integration with Markov-chain importance sampling is a common method.

#### Monte Carlo Integration

Typical numerical integration techniques, such as Trapezoidal/Simpson's rule are not feasible for thermodynamic integrals, as too many function evaluations must be made. Rather, Monte Carlo integration is used. The basic theorem is that

$$\int_V f(\vec{x}) dD x \approx V \langle f \rangle \pm V \sqrt{\frac{\langle f^2 \rangle - \langle f \rangle^2}{N}}$$

where

$$\langle f \rangle = \frac{1}{N} \sum_i f(\vec{x}_i) \quad \langle f^2 \rangle = \frac{1}{N} \sum_i f(\vec{x}_i)^2$$

and  $V$  is the volume being integrated over.  $N$  points are chosen independently and randomly with uniform probability distribution through this volume to obtain the average and variance of the function. This method derives its validity through the Law of Large Numbers and the Central Limit Theorem.<sup>3</sup>

---

<sup>3</sup>For 1D integrals, this is fairly inefficient, as  $N \sim 1 \times 10^5$  samples must be taken for good accuracy



This method works well if the function is relatively flat; if it is sharply peaked, importance sampling can be used to greatly improve the efficiency of the integration. Recall the riemann sum of integration:

$$\int_a^b f(x) dx \approx \frac{b-a}{N} \sum_j^N f(x_j)$$

Instead, we choose a  $g(x) > 0$  with  $\int_a^b g dx = 1$  such that  $f = hg$ . Then,

$$\int_a^b f(x) dx = \int_a^b h(x)g(x) \approx \frac{b-a}{N} \sum_j^N h(x_j)$$

where  $x_j$  are chosen with probability density  $g(x_j)$  and have the observable  $h(x_j)$ .

## Markov Chains

The probability densities are chosen via a stationary stochastic process, and the stationary stochastic process is determined via Markov Chains.

A *discrete markov chain* is a stochastic process withc generates a sequence of states with probabilities depending on the current state. A state cah be in one of  $R$  states,  $s_i$ . If a state is in state  $s_i$ , the probability of moving to state  $s_j$  is given by the transition probability  $p_{ij}$ . These transition probabilities form a matrix  $P$  known as a *transition matrix* or *Markov matrix*. This generalizes naturally to continuous systems, such as our integration.

A probability vector  $w$  is called *stationary*, *invariant*, or a *fixed-point* if

$$w^\top = w^\top P$$

The probability vector is a description of the system. Repeated steps (or *thermalization*)

$$w^\top = w^\top P^n$$

converges to the fixed point, whereupon it reaches equilibrium. The sufficient condition for a Markov chain to have such a fixed point is *detailed balance*:

$$w_i P_{ij} = w_j P_{ji}$$

We want the the fixed point of the Markov Chain to be  $P_{eq}(q)$ .

To find this probability density, we need to sample. For a canonical ensemble, we generate  $q$  with the probability distribution

$$P_{eq}(q) = \frac{e^{-\beta V(q)}}{\int dq' e^{-\beta V(q')}}$$

If the ensemble weight is real and positive, we can use this as a probability interpretation for the Monte Carlo integration. The simplest way to construct the Markov transition matrix  $P(\tilde{q} \leftarrow q)$  is via the *Metropolis-Hastings* method.<sup>4</sup>

---

<sup>4</sup>The probability distribution  $P_{eq}$  is not the same as the transition matrix  $P(\tilde{q} \leftarrow q)$

### Metropolis-Hastings Method

A proposal density  $R(\tilde{q} \leftarrow q)$  which is normalized, can be evaluated for all  $\tilde{q}, q$  and can be easily sampled is chosen—note, that  $R$  does not have to have any relationship to  $P_{eq}$ .

We use the matrix  $R$  to propose a new  $\tilde{q}$ . The new value  $\tilde{q}$  is accepted with probability

$$P_{acc} = \min \left( 1, \frac{R(q \leftarrow \tilde{q}) P_{eq}(\tilde{q})}{R(\tilde{q} \leftarrow q) P_{eq}(q)} \right)$$

If the proposal density is reversible

$$R(\tilde{q} \leftarrow q) = R(q \leftarrow \tilde{q})$$

the acceptance becomes

$$P_{acc} = \min \left( 1, \frac{P_{eq}(\tilde{q})}{P_{eq}(q)} \right)$$

thus, the integral need not be evaluated and the acceptance condition is related to a ratio of boltzmann factors. To adequately travel through phase space, we want the acceptance of a new state to be about 50%; this can be adjusted by modifying  $R$ .

This algorithm satisfies detailed balance, so it will eventually reach a fixed point.

## I have no idea what happened here

Typical computations are for estimating the expectations for potential  $\langle V \rangle, \langle V^2 \rangle$  to obtain

$$E = \frac{3N}{2\beta} + \langle V \rangle$$

$$c_V = \frac{3}{2}k_B + \frac{k_B\beta^2}{N} \left( \langle V^2 \rangle - \langle V \rangle^2 \right)$$

## 18.5 Factorization of the Partition Function

If particles don't interact with each other, but just an external potential, the partition function factorizes naturally. Take

$$H(q, p) = \sum_{j=1}^N \frac{p_j^2}{2m} + \sum_{j=1}^N \phi(r_j)$$

such a hamiltonian is satisfied if  $\phi$  is the mean field of all of the particles, for example. The partition function then naturally becomes

$$Z = \frac{1}{h^{3N} N!} (2\pi m k_B T)^{3N/2} \left( \int d^3r \exp[-\beta\phi(r)] \right)^N \quad (18.15)$$

Such an integral can always be evaluated (even if only numerically) so we can consider such problems to be solved.

### 18.5.1 Simple Harmonic Oscillator

A collection of uncoupled simple harmonic oscillators may be solved using factorization. Take the hamiltonian to be

$$H = \sum_{j=1}^N \frac{p_j^2}{2m_j} + \frac{1}{2} K_j x_j^2$$

Because the particles are localized, particle exchange is not an issue, so we do not need to divide by  $N!$  to account for degeneracy. Thus, the partition function may be easily seen to take the form

$$Z = \frac{1}{h^N} \prod_{j=1}^N (2\pi m_j k_B T)^{1/2} (2\pi k_B T / K_j)^{1/2} \quad (18.16)$$

Taking

$$\omega_j = \sqrt{\frac{K_j}{m_j}}$$

we can simplify this as

$$Z = \prod_{j=1}^N \frac{1}{\beta \hbar \omega_j}$$

### 18.5.2 Lennard-Jones Potential

Another common potential is the Lennard-Jones potential:

$$\phi \sim \left(\frac{r_0}{r}\right)^{12} - \left(\frac{r_0}{r}\right)^6 \quad (18.17)$$

The 6th power term comes from the interaction of two induced dipoles ( $(r^3)^2$  from temporary dipoles) while the 12th power term comes from the Pauli Exclusion repulsion force.

### 18.5.3 Mayer Cluster Expansion

For a pairwise potential, the partition function becomes

$$Z = \frac{1}{h^{3N} N!} \left(\frac{2\pi m}{\beta}\right)^{3N/2} \int \prod_{i=1}^N (d^3 r_i) \exp \left[ -\beta \sum_{j < k} \phi(|r_j - r_k|) \right] \quad (18.18)$$

We cannot expand the exponential, as the factor is not small. Instead, we expand in terms of

$$f(r) \equiv e^{-\beta \phi(r)} - 1 \quad (18.19)$$

in the *Mayer Cluster Expansion*. Let

$$f_{ij} = f(|r_i - r_j|)$$

Then, the partition function becomes

$$Z = \frac{1}{h^{3N} N!} \left( \frac{2\pi m}{\beta} \right)^{3N/2} \int \prod_{i=1}^N (d^3 r_i) \prod_{j < k} (1 + f_{jk})$$

Using logarithms and exponentials, we can further rewrite as

$$Z = \frac{1}{h^{3N} N!} \left( \frac{2\pi m}{\beta} \right)^{3N/2} \int \prod_{i=1}^N (d^3 r_i) \exp \left[ \sum_{j < k} \ln(1 + f_{jk}) \right]$$

We define

$$L = \sum_{j < k} \ln(1 + f_{jk}) \quad F_n = \sum_{j < k} f_{jk}^n$$

so we can expand  $L$  as

$$L = F_1 - \frac{1}{2} F_2 + \frac{1}{3} F_3 - \frac{1}{4} F_4 + \dots$$

Expanding the exponential,

$$\exp[L] = 1 + F_1 + \frac{1}{2} (F_1^2 - F_2) + \frac{1}{6} (F_1^3 - 3F_1 F_2 + 2F_3) + \dots$$

Considering only the first two terms, the first term yields  $V^N$  in the after integration—the same as for an ideal gas. The second term involves summing over  $\frac{1}{2}N(N-1)$  pairs, yielding

$$\int d^3 r \sum_{j < k} f_{jk} = \frac{1}{2} N(N-1) V^{N-2} \int d^3 r_1 d^3 r_2 f(r_{12})$$

Changing variables to  $R = \frac{1}{2}(r_1 + r_2)$  and  $r = r_1 - r_2$ , we can further simplify to

$$\frac{1}{2} N(N-1) V^{N-1} \int d^3 r f(r)$$

Because  $f$  only acts on atomic-scale distances, transforming the limits of integration provides a negligible effect. For large  $N$ ,  $N-1 \approx N$ , so we obtain

$$Z = \frac{1}{h^{3N} N!} \left( \frac{2\pi m}{\beta} \right)^{3N/2} V^N \left[ 1 + \frac{N^2}{2V} \int d^3 r f(r) \right] \quad (18.20)$$

We can then obtain the the free energy as

$$\begin{aligned} F &= -\frac{1}{\beta} \ln Z \\ &= -k_B T \left( X + N \ln V + \ln \left[ 1 + \frac{N^2}{2V} \int d^3 r f(r) + \dots \right] \right) \\ &\approx -k_B T \left[ X + N \ln V + \frac{N^2}{2V} \int d^3 r f(r) + \dots \right] \end{aligned}$$

Differentiating, we can obtain the *Virial Expansion*,

$$P = \frac{Nk_B T}{V} \left( 1 - \frac{N}{V} \int d^3r f(r) + \cdots \right)$$

Thus, using the cluster expansion, we can obtain the *Virial Coefficients*,  $B_j(T)$  via

$$\frac{P}{k_B T} = \frac{N}{V} + B_2(T) \frac{N^2}{V^2} + B_3(T) \frac{N^3}{V^3} + \cdots$$

The clusters can be viewed as graphs; these graphs are necessary to evaluate higher order Mayer cluster expansion terms. it turns out that if we consider only connected clusters, we can obtain the Grand Canonical Partition Function as the exponential of all connected diagrams.



# Chapter 19

## Canonical Ensembles

### 19.1 Grand Canonical Ensemble

The grand canonical ensemble is a system in contact with both thermal and particle reservoirs, so the system is at a fixed temperature  $T$  and chemical potential  $\mu$ . The relevant thermodynamic potential is

$$U[T, \mu] = U - TS - \mu N \quad (19.1)$$

which is known as the *grand potential*.<sup>1</sup> For extensive systems, we can apply the Euler relation and obtain

$$U[T, \mu] = -PV$$

#### 19.1.1 Probability Distribution

The system+reservoir is isolated from the rest of the universe, so

$$E_T = E + E_R \quad N_T = N + N_R$$

Recall that

$$P(E, N) = \frac{\Omega(E, N)\Omega_R(E_T - E, N_T - N)}{\Omega_T(E_T, N_T)}$$

where

$$\Omega_\alpha(E_\alpha, N_\alpha) = \frac{1}{h^{3N_\alpha} N_\alpha!} \iint \delta(E_\alpha - H_\alpha) dq_\alpha dp_\alpha$$

Taking the logarithm of the probability, and noting that the reservoir is much larger than the system, we can write:

$$\ln P(E, N) \approx \ln \Omega(E, N) + \ln \Omega_R(E_T, N_T) - E \left( \frac{\partial}{\partial E} \ln \Omega_R(E_T, N_T) \right)$$

---

<sup>1</sup>This is sometimes also known as the Landau potential. Confusingly, this is often denoted  $\Omega$ , which conflicts with the microcanonical partition function. Luckily, we can distinguish via context; the Grand Potential is a thermodynamic potential, whereas the Microcanonical Partition Function is a statistical mechanics partition function.

$$- N \left( \frac{\partial}{\partial N} \ln \Omega_R(E_T, N_T) \right) - \ln \Omega_T(E_T, N_T) + \dots$$

Once again recognizing  $\Omega = \frac{S}{k_B}$ , we can rewrite the derivatives to obtain

$$\ln P(E, N) = \beta - \beta\mu + \ln \Omega(E, N) + \ln \Omega_R(E_T, N_T) - \ln \Omega_T(E_T, N_T)$$

We define the quantity

$$\ln \mathcal{Z} = \ln \Omega_T(E_T, N_T) - \ln \Omega_T(E_T, N_T)$$

or

$$P(E, N) = \frac{1}{\mathcal{Z}} \Omega(E, N) e^{-\beta E} e^{\beta\mu N}$$

thus we obtain the grand canonical partition function

$$\mathcal{Z}(T, V, \mu) = \int_{N=0}^{\infty} \int_0^{\infty} dE \Omega(E, V, N) \exp[-\beta E + \beta\mu N] \quad (19.2)$$

We can see that the grand canonical partition function is related to the partition function via

$$\mathcal{Z}(T, V, \mu) = \sum_{N=0}^{\infty} Z(T, V, N) e^{\beta\mu N} \quad (19.3)$$

so  $\mathcal{Z}$  is a discrete Laplace transform of  $Z$ . In terms of the hamiltonian, we thus have

$$\mathcal{Z}(T, V, \mu) = \sum_{N=0}^{\infty} \int dq^{(N)} \int dp^{(N)} e^{-\beta H^{(N)}(q,p)} e^{\beta\mu N} \quad (19.4)$$

The term  $z = e^{\beta\mu}$  is known as the *fugacity*, so we may rewrite the partition function as

$$\mathcal{Z}(T, V, z) = \sum_{N_i} Z(T, V, N_i) z^{N_i} \quad (19.5)$$

Similar to the Canonical Partition Function, all of the thermodynamic information may be obtained from the Grand Canonical Partition Function. In a similar argument, we see that

$$\ln \mathcal{Z} = -\beta U[T, \mu] \quad (19.6)$$

Using this, we may obtain all other thermodynamic information

## 19.2 Classical Ideal Gas

Calculating the parittion function for the classical ideal gas, we can obtain

$$\mathcal{Z} = \sum_{N=0}^{\infty} \frac{1}{N!} \left( \frac{(2\pi m k_B T)^{3/2} V e^{\beta\mu}}{h^3} \right)^N = \exp \left[ \frac{(2\pi m k_B T)^{3/2} V e^{\beta\mu}}{h^3} \right]$$

We may then express the grand canonical potential as

$$-\beta U[T, \mu] = -PV = \ln \mathcal{Z}$$

so we see that

$$P = k_B T (2\pi m k_B T)^{3/2} h^{-3} e^{\beta\mu}$$



## 19.3 Other Classical Ensembles

It turns out that every thermodynamic potential corresponds to an ensemble in statistical mechanics. The legendre transforms between potentials corresponds to a discrete Laplace transform between ensembles. For each ensemble, the logarithm can be used to recover the potential.

### 19.3.1 Gibbs Free Energy

We define Gibbs partition function as

$$Z_G(T, P, N) = \int_0^\infty dV \int \frac{d^{3N}p \, d^{3N}q}{h^{3N} N!} e^{-\beta(H+PV)} \quad (19.7)$$

This can be rewritten as

$$Z_G = \int_0^\infty dV Z e^{-\beta PV}$$

We can obtain the Gibbs Free Energy in the expected way:

$$G(T, P, N) = -k_B T \ln Z_G(T, P, N)$$

#### Legendre Transform Alt. Notation

An alternative (more rigorous) notation of the legendre transform is given as:

$$G(T, P, N) = \min_V \{F(T, V, N) + PV\}$$



# Chapter 20

## Entropy Revisited

The Boltzmann definition of the Entropy only applies to the microcanonical ensemble; for the canonical and grand canonical ensemble, we must change the definition. Contrary to the Boltzmann entropy, which posits that each microstate are equally probable, in the (grand) canonical ensemble, the microstates are not so. Instead, we use the Gibbs entropy, which is of the form

$$S = -k_B \sum_j p_j \ln p_j \quad (20.1)$$

where  $p_j$  is the probability of energy  $E_j$ . However, calculating this quantity is tedious; rather, we prefer to calculate it from the partition functions.

### 20.1 From the Partition Functions

Recall that

$$S = - \left( \frac{\partial F}{\partial T} \right)_{V,N}$$

so, for the canonical ensemble, we have

$$S = - \frac{\partial \beta}{\partial T} \frac{\partial}{\partial \beta} \left( - \frac{\ln Z}{\beta} \right)$$

or

$$S = k_B \left( \ln Z - \beta \frac{\partial \ln Z}{\partial \beta} \right) \quad (20.2)$$

Similarly, for the grand canonical ensemble, we have the relation

$$S = - \left( \frac{\partial U[T, \mu]}{\partial T} \right)_{V, \mu} \quad (20.3)$$

so

$$S = k_B \left( \ln \mathcal{Z} - \beta \frac{\partial \ln \mathcal{Z}}{\partial \beta} \right) \quad (20.4)$$

## 20.2 Massieu Functions

In quantum mechanics, there are some instances where entropy is not monotonically increasing wrt energy, which allows for negative temperature. In such cases, obtaining  $U$  from  $S$  is problematic. Rather, entropy can be defined using Massieu Functions.

# Chapter 21

## Irreversibility in Thermodynamics

Recall that by the second law of thermodynamics, that if a constraint on a system is released, the entropy of the system cannot decrease. Such a process is irreversible, since the opposite transition would require the entropy to decrease, which is forbidden. However, most underlying microscopic physics *is* reversible. It is interesting to examine why macroscopic physics can be irreversible when microscopic physics is reversible.

### 21.0.1 A Cup of Coffee

For example, take a cup of coffee. At a time  $t = 0$ , the coffee is hot, but at a later time  $t_F$  the coffee is at the same temperature as the room. In a Gedankenexperiment, at time  $t_F$ , we reverse the velocities of all particles in the coffee, cup, room, etc. Thus, because newton's equations are reversible, the coffee should go back to being hot at  $t = 2t_F$ , but this is never observed.

### Phase Space Considerations

We can explain this using Liouville's theorem and fluid dynamics. We can view the hot coffee state as a drop of concentrated dye in a cup of water, or as a cluster of points in phase space. As time evolves, the ink moves through cup with constant local density, but the ink deforms from a normal drop and grows numerous tendrils and spreads throughout phase space. From a macroscopic viewpoint, the ink appears to mix almost homogeneously into the water. The Gibbs Entropy is meant to describe macroscopic physics; rather than consider the microscopic regions of ink and water, we treat it as an average ink density. In this analogy, the gibbs entropy is a coarse grained probability distribution

$$S = -k_B \int dq \, dp \, \bar{P} \ln \bar{P}$$

where  $\bar{P}$  is the average of  $P$  over some local region  $\Lambda$  which is macroscopically small but microscopically large. By Jensen's inequality, which states for a concave function  $f(x)$  that

$$f(\langle x \rangle) \geq \langle f(x) \rangle \quad (21.1)$$

we see that the integrand  $-P \ln P$  is a concave function, so the coarse-grained probability is never less than the true entropy using microscopic probability. The increase in entropy is linked to our (lack of) knowledge about the system, rather than the time-dependence of the system.

Returning to the cup of coffee, there only exists a single microstate in which the velocities reverse, that the cup of coffee may heat up again; the overwhelming number of microstates do not have this occur, and so statistically the coffee will not heat up. Further, as the system is a chaotic system, even if one velocity doesn't get reversed, the time evolution will be drastically different.

### 21.0.2 Poincaré recurrence theorem

The Poincaré recurrence theorem states that any isolated classical system must exhibit quasi-periodic behaviour—systems must repeatedly return to points in phase space that are arbitrarily close to its initial point. We reconcile this with irreversibility as the periodicity of each particle is different, so the overall period of the total system is incredibly large, even compared to the age of the universe.

Furthermore, we cannot predict recurrence time with uncertainty that is less than the age of the universe due to the lack of knowledge of all exact initial velocities, and the extreme sensitivity of Poincaré recurrence time to small changes in the initial velocities.

## Part IV

# Quantum Statistical Mechanics





# Chapter 22

## Quantum Mechanics

Rather than an in-depth review of quantum mechanics, we instead state some remarks comparing quantum mechanics to classical mechanics in the context of statistical mechanics. We will use a simple view, in 1D, non relativistic, and no spin. These are summarized in the table on the following page. Here are some important remarks about the table.

Pure states can be characterized as states of *minimal entropy*, and in classical mechanics represent a single point in phase space. However, due to the Heisenberg uncertainty principle, this does not translate well to quantum mechanics. Rather, in quantum mechanics, a minimum entropy state must be a “minimized blob” in phase space.

Eigenstates can be characterized as states for which repeated ideal measurements of the corresponding observable don’t scatter—there is no variance. We have shown that  $\sigma = 0$  implies eigenvectors; we will show the converse now. Assume  $AW = \alpha W$ . Then,

$$\sigma^2 = \langle A^2 \rangle - \langle A \rangle^2 = \text{tr}(A^2 W) - \text{tr}(AW)^2 = \text{tr}(\alpha^2 W) - \text{tr}(\alpha W)^2 = 0$$

---

<sup>1</sup> $L^2(\mathbb{R})$  is the space of square-integrable functions defined on the real line.

<sup>2</sup>A Liouville Space is an example of an “Algebra.”

<sup>3</sup>The operator  $\int_{\Gamma} \frac{dp dq}{h}$  is often called the trace over phase space, due to the analogy with quantum mechanics.

<sup>4</sup>A Lie product is a anti-symmetric bilinear operator satisfying the Jacobi Identity:  $\{a, \{b, c\}\} + \{b, \{c, a\}\} + \{c, \{a, b\}\} = 0$ . The vector cross product in  $\mathbb{R}^3$  is such a Lie product.

<sup>5</sup>Occasionally, the commutator is given the prefactor  $\frac{1}{i\hbar}$ . If  $A$  and  $B$  are hermitian, then the commutator is purely imaginary; further, most commutators have a factor of  $\hbar$ , which goes to zero in the classical limit. For example, this definition would make the canonical commutation relation  $[\hat{x}, \hat{p}] = 1$  rather than  $i\hbar$ .

<sup>6</sup>The notation  $W \geq 0$  is used to denote the fact that  $W$  must be positive semi-definite, or that all eigenvalues are non-negative. It is more useful to have the notion that the Statistical Operator is the state rather than the waveket. This is because mixed states can be represented by statistical operators, while wavekets can only denote pure states.

<sup>7</sup>A pure state is contrasted to a mixed state where the former is a well defined state, while the second is a probability distribution of states (such as a macrostate in Stat Mech, or the more familiar notion in matrix mechanics)

<sup>8</sup>The classical entropy is motivated by Shannon’s notion of information entropy.

<sup>9</sup>The Liouville equation defines the equation of motion for a probability density through phase space in classical mechanics. The quantum analogue is the von Neumann equation. The analogy between these two equations is another motivation for the factor of  $\frac{1}{i\hbar}$  in the commutator

Table 22.1: Table 1: Comparison of Classical and Quantum Mechanics

Concept	Classical Mechanics	Quantum Mechanics
“arena”	Phase Space $\Gamma \simeq \mathbb{R}^2$	Hilbert Space <sup>1</sup> $\mathcal{H} \simeq L^2(\mathbb{R})$
Canonical Variables	real coordinates $p, q$	hermitian operators $\hat{P}, \hat{Q}$
Elements of Liouville Space <sup>2</sup>	Complex functions on $\Gamma$ $a(p, q)$	operators on $\mathcal{H}$ $A(P, Q)$
Scalar Products <sup>3</sup>	$\langle a, b \rangle = \int_{\Gamma} \frac{dp dq}{h} a^* b$	$\langle A, B \rangle = \text{tr}(A^\dagger B)$
Lie Product <sup>4</sup>	Poisson Bracket $\{a, b\} = \frac{\partial a}{\partial q} \frac{\partial b}{\partial p} - \frac{\partial a}{\partial p} \frac{\partial b}{\partial q}$	Commutator <sup>5</sup> $[A, B] = AB - BA$
Observables	Real functions on $\Gamma$	Hermitian operators on $\mathcal{H}$
States	Probability densities on $\Gamma$ $*w = w \geq 0, \langle w, 1 \rangle = 1$	Statistical Operators <sup>6</sup> on $\mathcal{H}$ $W^\dagger = W \geq 0, \langle W, 1 \rangle = 1$
Pure States <sup>7</sup>	A point in phase space $\Gamma$ $w(p, q) = h\delta(p - p_0)\delta(q - q_0)$	Idempotent operators $W = W^2 \iff W =  \psi\rangle\langle\psi $
Eigenstate of an Observable	$w(p, q) = \rho(p, q)\delta[a(p, q) - \alpha]$	$AW = \alpha W, \alpha \in \mathbb{R}$
Expectation Values of an Observable in State $a$	$\langle a \rangle = \langle w, a \rangle$	$\langle A \rangle = \langle W, A \rangle$
Property/Event	Characteristic functions on a subset $\mathcal{E} \subset \Gamma$ $\chi_{\mathcal{E}}(p, q) = \begin{cases} 1 & (p, q) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases}$	Projection $E$ onto a subspace $\mathcal{E} \subset \mathcal{H}$
Probability of Event	$\langle \chi_E \rangle$	$\langle E \rangle$
Entropy of State <sup>8</sup>	$-k_B \langle w, \ln w \rangle$ $= -k_B \int \frac{dp dq}{h} w \ln w$	$-k_B \langle W, \ln W \rangle$
Liouville Equation <sup>9</sup>	$\frac{\partial w}{\partial t} = \{w, h\}$ for hamiltonian $h$	$\frac{\partial W}{\partial t} = -\frac{1}{i\hbar} [W, H]$

Similarly, pure eigenstates can be characterized by normalized eigenvectors such  $|\psi\rangle$  such that  $W = |\psi\rangle\langle\psi|$  and  $\langle\psi|\psi\rangle = 1$

$$\begin{aligned} AW &= (A|\psi\rangle)\langle\psi| = (\alpha|\psi\rangle)\langle\psi| = \alpha W \\ W^2 &= |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = \langle\psi|\psi\rangle|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| \end{aligned}$$

Mixed states are *convex combinations* of states  $W_j$  with coefficients  $p_j$  such that  $\sum_j p_j = 1$ . The mixed state  $W$  is then given

$$W = \sum_j p_j W_j$$

Trivially, for real coefficients, the linearity of the adjoint and trace guarantee that  $W$  is trace 1 and self-adjoint. The positive-semidefiniteness follows from the positive real coefficients on the sum of positive-semidefinite matrices. If all of the  $W_j$  are pure and belong to the same eigenspace  $AW_j = \alpha W_j$ , then  $W$  is not pure, but is still an eigenstate of  $A$ . Thus, *eigenstates need not be pure states*.

Expectation values are “weighted means” of a stochastic observable. Given a state  $W$ , with eigenvectors  $|\psi_n\rangle$  such that  $\langle\psi_n|\psi_m\rangle = \delta_{nm}$  and  $\sum_n |\psi_n\rangle\langle\psi_n| = \mathbb{1}$ , then, we can write

$$W = \sum_n p_n |\psi_n\rangle\langle\psi_n|$$

for suitable coefficients  $p_n$  such that  $0 \leq p_n \leq 1$  and  $\sum_n p_n = 1$ . This is the *spectral theorem* for a general quantum state. This spectral decomposition is useful in computing expectation values:

$$\begin{aligned} \langle A \rangle &= \text{tr}(A^\dagger W) = \text{tr}(W^\dagger A) = \text{tr}(WA) \\ &= \text{tr} \left( \sum_n p_n |\psi_n\rangle\langle\psi_n| A \right) \\ &= \sum_{nm} \langle\psi_m| p_n |\psi_n\rangle \langle\psi_n| A |\psi_m\rangle \\ &= \sum_n p_n \langle\psi_n| A |\psi_n\rangle \end{aligned}$$

The term  $p_n$  in the sum is what Deserno calls “subjective ignorance,” or that we haven’t prepared our state carefully enough, which leads to error. The term  $\langle\psi_n| A |\psi_n\rangle$  is what he calls “objective indeterminacy,” which is that some states cannot be well defined given perfect conditions. In other words, the former is limitations due to our preparation, while the latter is limitations due to the nature of physics. Thus, there are two sources for stochasticity in quantum mechanics.

The entropy of a state is then a measure of both our subjective ignorance and objective indeterminacy. This leads to a difference between classical and quantum mechanics. In classical mechanics, the entropy is bounded by<sup>10</sup>

$$-\infty \leq S_{cl} \leq k_B \ln \left( \int_{\Gamma} \frac{dq dp}{h} \right) = +\infty$$

---

<sup>10</sup>in  $\Gamma \simeq \mathbb{R}^2$

However, in quantum mechanics, the entropy is bounded below<sup>11</sup>

$$0 \leq S_{qm} \leq k_B \ln(\dim \mathcal{H}) = +\infty$$

In the proof of the classical version, the entropy of a Gaussian can be used to show the unboundedness. We will show the proof of the quantum version. From the spectral decomposition, we can show that for  $W = \sum_n p_n |\psi_n\rangle\langle\psi_n|$

$$W \ln W = \sum_n p_n \ln p_n |\psi_n\rangle\langle\psi_n|$$

The entropy is then

$$S = -k_B \operatorname{tr}(W \ln W) = -k_B \sum_n p_n \ln p_n$$

The minimum entropy (maximal knowledge) state is that of a pure state:  $p_n = \delta_{n,n_0} \implies S = 0$ . The maximal entropy (minimal knowledge) state is given by a state where  $p_n = \frac{1}{\dim \mathcal{H}}$ . This gives the entropy as

$$S = -k_B \sum_{n=1}^N \frac{1}{N} \ln \frac{1}{N} = k_B \ln N = k_B \ln(\dim \mathcal{H}) \rightarrow \infty$$

An intermediate step,  $S = k_B \ln N$  should be familiar—we saw a similar statement of entropy earlier as  $S = k_B \ln \Omega$ .

### Functions of Operators

Under certain circumstances, (most importantly hermiticity with compactness<sup>12</sup>), an operator can be spectrally decomposed as

$$A = \sum_n \lambda_n |n\rangle\langle n|$$

Spectral decomposition is a very powerful property of an operator. For example, if we can define a function of an operator by using its Taylor series. To see why, when an operator is spectrally decomposed, it is trivial to take powers:

$$A^m = \sum_n \lambda_n^m |n\rangle\langle n|$$

From this, we use the Taylor series to show

$$f(A) = \sum_n f(\lambda_n) |n\rangle\langle n|$$

<sup>11</sup>in  $\mathcal{H} \simeq L^2(\mathbb{R})$

<sup>12</sup>a linear operator  $A$  on  $\mathcal{H}$  is considered compact if the closure of the image of any bounded subset of  $H$  is compact.

# Chapter 23

## Quantum Canonical Ensemble

By analogy to the classical case, we can reasonably guess that the quantum canonical state can be defined as follows:

$$W = \frac{1}{Z} e^{-\beta H} \quad (23.1)$$

As  $W$  is a function of the Hamiltonian operator  $H$ , it too is an operator. Once again, the factor  $Z$  is the partition function, chosen to normalize  $W$ . As such, we have that

$$Z = \text{tr}(e^{-\beta H}) \quad (23.2)$$

Subtleties arise in the trace due to symmetry constraints. We will revisit this idea later.

Because the Hamiltonian is self-adjoint (and hopefully compact), it has real eigenvalues and its eigenvectors form an orthonormal basis over the Hilbert space  $\mathcal{H}$ . We will denote this basis via

$$H|n\rangle = E_n|n\rangle \quad (23.3)$$

or equivalently, we should be able to spectrally decompose the Hamiltonian as

$$H = \sum_n E_n |n\rangle \langle n|$$

Since the canonical state is a function of  $H$ , it is diagonalizable in the same basis. Trivially, we can then compute the partition function as

$$Z = \text{tr}(e^{-\beta H}) = \sum_n \langle n | e^{-\beta H} | n \rangle = \sum_n e^{-\beta E_n}$$

This equation is deceptively simple. When we actually compute it for certain systems, it may well be harder than integrating a (grand)canonical partition function for a comparable classical system, as integration has more flexibility than sums. Further, we will often have degenerate energy levels. In this case, we can sum over energy *levels*, keeping in mind the degeneracy of those levels. Thus,

$$Z = \sum_n e^{-\beta E_n} = \sum_\ell \Omega(\ell) e^{-\beta E_\ell} \quad (23.4)$$

where the index  $n$  considers eigenstates,  $\ell$  considers energy levels, and  $\Omega(\ell)$  is the degeneracy of level  $\ell$ .

### Thermal Averages

From the quantum canonical state, we can then compute thermal averages:

$$\langle A \rangle = \langle W, A \rangle = \text{tr}(WA) = \text{tr}(AW) = \sum_n \langle n|AW|n \rangle$$

Because we chose the trace to be calculated along an eigenbasis of  $W$ , we can simplify this sum as

$$\langle A \rangle = \frac{1}{Z} \sum_n e^{-\beta E_n} \langle n|A|n \rangle \equiv \sum_n P_n \langle n|A|n \rangle \quad (23.5)$$

We can consider the term  $P_n = e^{-\beta E_n}/Z$  as, once again, a measure of our subjective ignorance. If we take  $A = H$ , then, we trivially have

$$E = \langle H \rangle = \sum_n P_n E_n$$

## 23.1 Entropy

We argue that the free energy can be represented in a manner similar to that of the classical canonical partition function as

$$F = -k_B T \ln Z \quad (23.6)$$

We motivate this by looking at the quantum entropy

$$S = -k_B \text{tr}(W \ln W) \quad (23.7)$$

which we also have not shown to be true (but should be compared to information entropy). Computing  $S$ , we obtain

$$\begin{aligned} S &= -k_B \sum_n \langle n|W \ln W|n \rangle \\ &= -k_B \sum_n \langle n| \frac{e^{-\beta E_n}}{Z} \ln \frac{e^{-\beta E_n}}{Z} |n \rangle \\ &= -k_B \sum_n \frac{e^{\beta E_n}}{Z} (-\beta E_n - \ln Z) \\ &= \frac{1}{T} E + k_B \ln Z \\ TS &= E + k_B \ln Z \end{aligned}$$

$$\begin{aligned} -k_B \ln Z &= E - TS \\ &= F \end{aligned}$$

Note that we have only shown that these arguments hold for the canonical state; however, the formula for entropy applies for *any* state; this is due to analogy from Shannon entropy from information theory.

### 23.1.1 Low Temperature Limit

As temperature decreases, we argue we can write

$$Z = \sum_{\ell} \Omega(\ell) e^{-\beta E_{\ell}} = \Omega(0) e^{-\beta E_0} \left[ 1 + \sum_{\ell > 0} \frac{\Omega(\ell)}{\Omega(0)} e^{-\beta(E_{\ell} - E_0)} \right]$$

We argue that the excited states ( $\ell > 0$ ) exponentially go to zero, leaving only the ground state contribution. Thus, for sufficiently low  $T$ , the partition function can be written

$$Z(T) \approx \Omega(0) e^{-\beta E_0}$$

Thus, the probability of being in any one of the ground states approaches a constant  $\frac{1}{\Omega(0)}$ , while the probability of being in an excited state is given

$$P_{\ell \neq 0} = \frac{\Omega(\ell)}{\Omega(0)} e^{-\beta(E_{\ell} - E_0)} \rightarrow 0$$

Because we now know the probabilities of every state, we can easily calculate the entropy:

$$\lim_{t \rightarrow 0} S = -k_B \sum_n \frac{1}{\Omega(0)} \ln \frac{1}{\Omega(0)} = k_B \ln \Omega(0)$$

Using this result, we see that in the thermodynamic limit<sup>1</sup>,

$$\lim_{N \rightarrow \infty} \lim_{T \rightarrow 0^+} \frac{S(T)}{N} = \lim_{N \rightarrow \infty} \frac{k_B \ln \Omega(0, N)}{N} = 0$$

which holds as long as the ground state isn't "exponentially degenerate", i.e.  $\Omega(0, N)$  does not scale with  $a^N$  for some  $a > 1$ . In such a case, we obtain the limit to be

$$\lim_{N \rightarrow \infty} \lim_{t \rightarrow 0^+} = \lim_{N \rightarrow \infty} \frac{k_B \ln a^N}{N} = k_B \ln a$$

Thus, the difference between the Nernst and Planck version of the 3<sup>rd</sup> law is whether we believe it to be possible for a system to have an exponentially degenerate ground state. However, it is possible to show that there are idealized models such that the ground state is exponentially degenerate; on the other hand, we do not know whether such models represent reality well.

---

<sup>1</sup>We cannot know if the limits can be interchanged, which provides some degree of questionability

### Planck in Classical Statistical Mechanics

As shown in the homework,

$$Z_{\text{quantum}} \leq Z_{\text{classical}}$$

it turns out the Golden-Thompson inequality gets closer and closer to equality as  $T \rightarrow \infty$ , or  $\beta \rightarrow 0$ . As such, we want to have in the large-temperature limit, we also want the inequality of partition functions to approach an equality. As such, we use Planck's constant as a scaling factor with units of action, as opposed to some other scaling factor.

Recall that Golden-Thompson shows us that

$$F_{\text{quantum}} = -k_B T \ln \text{tr} \left( e^{-\beta \left( \frac{P^2}{2m} + V(Q) \right)} \right) \geq -k_B T \ln \text{tr} \left( e^{-\beta \frac{P^2}{2m}} e^{-\beta V(Q)} \right) = F_{\text{classical}}$$

In Landau-Lifshitz v, §33, the following<sup>2</sup> is derived:

$$F_{\text{quantum}} = F_{\text{classical}} + \frac{1}{24} \frac{\hbar^2}{m(k_B T)^2} \langle (V'(q))^2 \rangle + \mathcal{O}(\hbar^3) \quad (23.8)$$

note, that the leading order correction term contains the thermal average of the square-forces on the system. Further, the correction term vanishes for  $T \rightarrow \infty$  and  $m \rightarrow \infty$ . The following motivates approximations such as the Born-Oppenheimer approximation, as the lighter particles experience larger quantum corrections than heavier particles.

## 23.2 Derivatives of Thermal Averages

Recall that we can compute the thermal average of an observable  $A$  as

$$\langle A \rangle = \sum_n \frac{e^{-\beta E_n}}{Z} \langle n|A|n \rangle \quad (23.9)$$

Interestingly, the quantum averaging term,  $\langle n|A|n \rangle$  has *no temperature dependence*; all of the temperature dependence comes in the leading, “post-quantum” averaging term. Thus,

$$\begin{aligned} \frac{\partial A}{\partial \beta} &= \sum_n \frac{\partial}{\partial \beta} \frac{e^{-\beta E_n}}{Z} \langle n|A|n \rangle \\ &= \sum_n \frac{(-E_n)e^{-\beta E_n} Z - e^{-\beta E_n} \frac{\partial Z}{\partial \beta}}{Z^2} \langle n|A|n \rangle \\ &= \sum_n \frac{e^{-\beta E_n}}{Z} \left[ -E_n - \frac{\partial \ln Z}{\partial \beta} \right] \langle n|A|n \rangle \end{aligned}$$

---

<sup>2</sup>This equation can be derived using the Zassenhaus formula



$$\begin{aligned}
&= - \sum_n E_n \frac{e^{-\beta E_n}}{Z} \langle n|A|n\rangle + \langle H \rangle \sum_n \frac{e^{-\beta E_n}}{Z} \langle n|A|n\rangle \\
&= - \sum_n \frac{e^{-\beta E_n}}{Z} \langle n|AH|n\rangle + \langle H \rangle \vec{A}
\end{aligned}$$

and so, we obtain the important result

$$\frac{\partial \langle A \rangle}{\partial \beta} = - \langle AH \rangle + \langle A \rangle \langle H \rangle = - \text{Cov}(A, H) \quad (23.10)$$

where Cov is the *covariance* of the two operators. This expression is also true classically.

Consider the specific heat

$$c_V = \frac{1}{N} \frac{\partial \langle H \rangle}{\partial T} = \frac{1}{N} \frac{\partial \beta}{\partial T} \frac{\partial \langle H \rangle}{\partial \beta} = \frac{1}{N k_B T^2} \text{Cov}(H, H) = \frac{\sigma_H^2}{N k_B T^2}$$

as, of course, the covariance of an operator with itself is the variance of the operator. Dividing by Boltzmann's constant, we obtain

$$\frac{c_V}{k_B} = \frac{1}{N} \left( \frac{\sigma_H}{k_B T} \right)^2 \quad (23.11)$$

This is an example of a *fluctuation-response theorem*, and is the same as the classical result. The quantity  $c_V$  is an example of a response, while  $\sigma_H$  is an example of a fluctuation. Fluctuation-response theorems show that responses result from fluctuations.

## 23.3 Merging fun!

## 23.4 Factorization of the Partition Function

Recall that in classical mechanics that the additivity of Hamiltonians lead to the factorization of the partition function:

$$H = \sum_j H_j \implies Z = \prod_j Z_j$$

Consider a collection of  $N$  particles. Assume we know the spectral decomposition of the each of the individual Hamiltonians.; that is we know

$$H_j |n_j\rangle = E_{n_j} |n_j\rangle$$

We consider the following state vector:

$$|n\rangle = \bigotimes_{j=1}^N |n_j\rangle$$

we can then extend the individual hamiltonians to act on this state vector:

$$\tilde{H}_j = \underbrace{\mathbb{1} \otimes \cdots \otimes \mathbb{1}}_{j-1 \text{ times}} \otimes H_j \otimes \underbrace{\mathbb{1} \otimes \cdots \otimes \mathbb{1}}_{N-j \text{ times}}$$

Then, we can write the overall hamiltonian as

$$H = \sum_{j=1}^N \tilde{H}_j |n\rangle = \left( \sum_{j=1}^N E_{n_j} \right) |n\rangle = E_n |n\rangle$$

so we have an eigenvalue problem in terms of the overall state ket  $|n\rangle$ .

Thus, we have the partition function as

$$\begin{aligned} Z &= \sum_{\{n_j\}} e^{-\beta \sum_{j=1}^N E_{n_j}} \\ &= \sum_{\{n_j\}} \prod_{j=1}^N e^{-\beta E_{n_j}} \\ &= \sum_{n_1} \cdots \sum_{n_N} e^{-\beta E_{n_1}} \cdots e^{-\beta E_{n_N}} \\ &= \left( \sum_{n_1} e^{-\beta E_{n_1}} \right) \cdots \left( \sum_{n_N} e^{-\beta E_{n_N}} \right) \\ &= \prod_{j=1}^N \sum_{n_j} e^{-\beta E_{n_j}} \\ &= \prod_{j=1}^N Z_j \end{aligned}$$

And so, under certain circumstances, we can factorize the partition function. However, there are two notes to this proof. First, note that we did not simply swap the sum and the product from line 2 to line 5; rather, the former has a sum over all possible energy configurations among all particles, while the latter is the sum over the energy states for a particular particle. Second, for fermions and bosons, we also need to account for (anti)symmetrizations of the state. This second point is a consequence of the indistinguishability of the particles and will be discussed later.

### 23.4.1 Two-Level System

Consider a system with a hamiltonian  $H = \varepsilon n$  with  $n = \{0, 1\}$ . For example, such a system could be represented by a particle in a magnetic field. We can then calculate the partition function for a single particle as

$$Z = \sum_{n=0}^1 e^{-\beta \varepsilon n} = 1 + e^{-\beta \varepsilon}$$

Now that we know the partition function, we now know the thermodynamics of the system. The free energy can be found

$$F = -k_B T \ln Z = -k_B T \ln(1 + e^{-\beta\varepsilon})$$

and the energy as

$$E = \frac{\partial \beta F}{\partial \beta} = \frac{\varepsilon}{1 + e^{+\beta\varepsilon}}$$

This expression makes sense; when  $T \rightarrow 0$ ,  $\beta \rightarrow \infty$  so the energy is 0. However, when  $T \rightarrow \infty$ ,  $\beta \rightarrow 0$ , the energy goes to  $\varepsilon/2$ . This is slightly unexpected, but can be rationalized by considering entropy. When  $T \rightarrow \infty$ , the entropic effect on the free energy dominates the thermodynamics, so the state attempts to maximise the entropy; this happens with equal occupation of the two states.

Similarly, we can determine the expectation of  $n$ :

$$\langle n \rangle = \sum_{n=0}^1 n \frac{e^{-\beta\varepsilon n}}{z} = \frac{1}{1 + e^{+\beta\varepsilon}} = \frac{E}{\varepsilon}$$

The fact that the denominator contains two positive signs has incredibly important consequences, as we will see later. If we plot  $\langle n \rangle$  as a function of the temperature, we obtain a characteristic temperature  $T^* = \varepsilon/k_B$  where the slope is the steepest, and asymptotes to 1/2 as  $T \rightarrow \infty$ . Unfortunately, there is no symmetry about the critical temperature. {diagram}

### 23.4.2 Simple Harmonic Oscillator

Consider the hamiltonian given

$$H = \frac{P^2}{2m} + \frac{1}{2}\omega^2 Q^2 = \hbar\omega \left( a^\dagger a + \frac{1}{2} \right) = \hbar\omega \left( \hat{n} + \frac{1}{2} \right)$$

The last form is convenient, as we already know the spectrum of  $\hat{n}$  as  $n \in \mathbb{N} \cup \{0\}$ . This gives us a very simple partition function to compute:

$$Z = \text{tr}(e^{-\beta H}) = \sum_{n=0}^{\infty} e^{-\beta\hbar\omega(n+1/2)} = e^{-\frac{1}{2}\beta\hbar\omega} \sum_{n=0}^{\infty} e^{-\beta\hbar\omega n} = e^{-\frac{1}{2}\beta\hbar\omega} \cdot \frac{1}{1 - e^{-\beta\hbar\omega}}$$

We can rewrite the partition as

$$Z = \frac{1}{2 \sinh \frac{\beta\hbar\omega}{2}} \tag{23.12}$$

In the limit  $T \rightarrow \infty, \beta \rightarrow 0$ , we obtain  $\sinh \frac{\beta\hbar\omega}{2} \rightarrow \frac{\beta\hbar\omega}{2}$ . Then, we obtain

$$Z \rightarrow \frac{k_B T}{\hbar\omega}$$

which is the classical result.

The free energy can of course be written as

$$F = k_B T \ln \left( 2 \sinh \frac{\beta \hbar \omega}{2} \right)$$

and the energy, after some calculation, is

$$E = \frac{\hbar \omega}{2} \tanh \frac{\beta \hbar \omega}{2} = \hbar \omega \left( \frac{1}{e^{\beta \hbar \omega} - 1} + \frac{1}{2} \right) \quad (23.13)$$

We recognize the expectation of the number operator is give

$$\langle n \rangle = \frac{1}{e^{\beta \hbar \omega} - 1} \quad (23.14)$$

notice how similar this occupancy looks to that of the two-level system, albeit with a  $-1$  instead of a  $+1$ . If we once again plot  $\langle n \rangle$  against  $T$ . We see as  $T \rightarrow \infty$ ,  $\beta \rightarrow 0$  from above, so there is no upper bound; however, it asymptotically reaches a linear behaviour at high temperature; the intercept of this asymptote is given

$$\frac{\hbar \omega}{k_B T} = \frac{1}{2}$$

# Chapter 24

## Black Body Radiation

### 24.1 Resonator Cavity

Recall Maxwell's equations:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0} \quad (24.1a) \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (24.1c)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (24.1b) \quad \nabla \times \mathbf{B} = \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \mathbf{J} \quad (24.1d)$$

Consider the application of these equations in a vacuum; that is  $\rho = 0$ ,  $\mathbf{J} = 0$ . We then have

$$\frac{1}{c^2} \ddot{\mathbf{E}} = \nabla \times \dot{\mathbf{E}} = -\nabla \times \nabla \times \mathbf{E} = -\nabla(\nabla \cdot \mathbf{E}) + \nabla^2 \mathbf{E} = \nabla^2 \mathbf{E}$$

Similarly, we find that

$$\frac{1}{c^2} \ddot{\mathbf{B}} = \nabla^2 \mathbf{B}$$

Thus, we see that both  $\mathbf{B}$  and  $\mathbf{E}$  satisfy the wave equation:

$$\nabla^2 \begin{Bmatrix} \mathbf{E} \\ \mathbf{B} \end{Bmatrix} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \begin{Bmatrix} \mathbf{E} \\ \mathbf{B} \end{Bmatrix} = 0$$

We will consider eigenmodes<sup>1</sup> inside a hollow metal cube  $[0, L]^3$ . Because the walls are an equipotential, we know that  $\mathbf{E}$  must be perpendicular to the walls. We guess that the electric field may be written:

$$\mathbf{E}(\mathbf{r}, t) = \begin{bmatrix} E_{x,0} \cos(k_x x) \sin(k_y y) \sin(k_z z) \\ E_{y,0} \sin(k_x x) \cos(k_y y) \sin(k_z z) \\ E_{z,0} \sin(k_x x) \sin(k_y y) \cos(k_z z) \end{bmatrix} \sin \omega t \quad (24.2)$$

and indeed, we can verify that this satisfies the wave equation:

$$\nabla^2 \mathbf{E} = -(k_x^2 + k_y^2 + k_z^2) \mathbf{E}$$

---

<sup>1</sup>We can view the wave equation as an eigenvalue equation for the D'Alembert operator  $\square \mathbf{E} = 0 \mathbf{E}$ .

$$\frac{\partial^2 \mathbf{E}}{\partial t^2} = -\omega^2 \mathbf{E}$$

so long as we have

$$\mathbf{k}^2 = \frac{\omega^2}{c^2}$$

Note that due to the symmetry of our wave, we need only that

$$\omega = c |\mathbf{k}|$$

This is the *dispersion relation* for electromagnetic waves.<sup>2</sup>

Further, our guess satisfies the boundary conditions, that  $\mathbf{E}$  must be perpendicular to the walls. Indeed, at  $x = 0$ , the term  $\sin(k_x x)$  kills off all components other than the  $\hat{x}$ , and similarly so for the other axial planes. However, when we consider  $x, y, z = L$ , we see that we must have

$$\sin(k_i L) = 0 \implies k_i L = \pi m_i$$

so we have a further constraint that

$$k_i = \frac{\pi m_i}{L} \quad m_i \in \mathbb{Z}$$

combining this with the wave equation, we obtain

$$\omega = \frac{\pi c}{L} \sqrt{m_x^2 + m_y^2 + m_z^2}$$

Once again, due to the symmetry of the wave, we need only consider  $m_i \in \mathbb{N} \setminus \{0\}$ . Thus,

$$\omega = \frac{\pi c m}{L} \tag{24.3}$$

where  $\mathbf{m} \in \mathbb{N}^3 \setminus \{0\}$  is called the *mode*. However, we need one more parameter to specify the state—the *polarization* of the wave,  $\sigma \in \{1, 2\}$ . It is interesting to note that the allowed states in the box are discretized.

### 24.1.1 Density of States

Consider the density of states in mode space:

$$D(\mathbf{m}) = 2 \sum_{\mathbf{n}} \delta(\mathbf{m} - \mathbf{n})$$

where the factor of 2 is included due to the polarization. We can extend this to a continuous variable, so

$$D(\mathbf{m}) = 2$$

---

<sup>2</sup>This can be rewritten  $E = \hbar \omega c |\mathbf{k}|$ . The dispersion relation tells us how wavepackets behave, as we it encodes both the group velocity and phase velocity.

We also want the DoS in frequency space; this can be obtained by

$$\begin{aligned}
 D(\omega) &= \int d^3\mathbf{m} D(\mathbf{m}) \delta\left(\omega - \frac{\pi c m}{2}\right) \\
 &= \frac{4\pi}{8} \int dm m^2 D(m) \delta\left(\omega - \frac{\pi c m}{2}\right) \\
 &= \frac{\pi}{2} \cdot 2 \left(\frac{L}{\pi c}\right)^3 \int d\tilde{m} \tilde{m}^2 \delta(\omega - \tilde{m}) \\
 &= \left(\frac{L}{\pi c}\right)^3 \pi \omega^2
 \end{aligned}$$

The factor of  $\pi\omega^2$  can be viewed as a multiplicity constant due to the portion of a spherical shell in the first octant:  $\frac{1}{8} \cdot 2 \cdot 4\pi\omega^2$ , keeping in mind the polarization factor 2.

## 24.2 Quantization

We notice that every eigenmode of the EM field is essentially a harmonic oscillator of frequency  $\omega$ ; following Planck, we consider each of these oscillators to be as a collection of quantum mechanical oscillators. The individual hamiltonian then becomes

$$\hat{H}_{\mathbf{m},\sigma} = \hbar\omega_{\mathbf{m}} \left( \hat{n}_{\mathbf{m},\sigma} + \frac{1}{2} \right) - \frac{1}{2}\hbar\omega_{\mathbf{m}}$$

where the second term is to neglect the ground-state energy in analogy to the classical expectation. And so, the total hamiltonian of the electromagnetic field in the box can be given

$$\hat{H} = \sum_{\mathbf{k},\sigma} \left[ \hbar c k \left( \hat{n}_{\mathbf{k},\sigma} + \frac{1}{2} \right) - \frac{1}{2} \hbar c k \right] \quad (24.4)$$

and so the EM field is equivalent to two quantum harmonic oscillators at every valid momentum space point. Note that our calculation has two different quantizations. First, is that the classical boundary conditions and their corresponding eigenmodes enforce a discretized, but not quantum, momentum space. However, the quantization of the energy levels of the harmonic oscillators is the imposition of a quantum mechanical effect. The energy expectation for a single oscillator, is of course,

$$\langle \varepsilon_{\mathbf{k},\sigma} \rangle = \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1}$$

Thus, the spectrally resolved energy density is

$$u_{\omega} = \frac{1}{V} D_{\omega}(\omega) \langle \varepsilon_{\omega} \rangle$$

where we divide by volume to obtain a density, and the remaining terms are the energy of a state times the density of that state. Substituting and simplifying, we obtain

$$u_{\omega} = \frac{\hbar\omega^3}{\pi^2 c^3} \cdot \frac{1}{e^{\beta\hbar\omega} - 1} \quad (24.5)$$

### 24.2.1 Energy Flux

If we poke a hole into the box and allow radiation to escape, we obtain an energy flux through the hole is

$$j_\omega = \frac{1}{4}cu_\omega$$

where of course,  $cu_\omega \sim v\rho = j$ . The prefactor of  $1/4$  arises due to the average projection over a given direction of a half sphere (particles going left will not exit to the right). Let the hole be pointing in the  $+\hat{z}$  direction. As photons exit the hole, they point in any direction. The flux along this axis must then be proportional to

$$\frac{\int_0^{\pi/2} d\theta \sin \theta \int_0^{2\pi} d\phi \cos \theta}{\int_0^{\pi/2} d\theta \sin \theta \int_0^{2\pi} d\phi} = \frac{1}{4}$$

Thus we obtain

$$j_\omega = \frac{\hbar}{4\pi^2 c^2} \cdot \frac{\omega^3}{e^{\beta\hbar\omega} + 1} \quad (24.6)$$

which is more commonly known as the planck law. At low  $\omega$ , this grows  $\sim \omega^2$ , which agrees with the classical density of states, but at high  $\omega$ , the quantum effects dominate, so it falls off with  $\sim \omega^2 e^{-\beta\hbar\omega}$  which allows the function to be integrable and resolve the UV catastrophe.

As shown in the HW, we can rewrite the density for wavelength space as

$$j_\lambda = \frac{2\pi\hbar c^2}{\lambda^5} \frac{1}{e^{\beta\hbar c/\lambda} + 1} \quad (24.7)$$

Fix  $x = \beta\hbar\omega$  and  $y = \frac{\beta\hbar c}{\lambda}$ . Then, we can write the distributions as

$$j_x = A_x \frac{x^3}{e^x - 1} \quad j_y = A_y \frac{y^5}{e^y + 1}$$

Finding the maximum, we obtain  $x_{\max} = 3 + W_0\left(-\frac{3}{e^3}\right) \approx 2.82144$ , where  $W_0$  is the Lambert  $W$  function, or the inverse<sup>3</sup> of  $f(x) = xe^x$ . The maximum flux can be found

$$j_{\omega_{\max}} = \frac{\hbar}{(2\pi c)^2} \frac{\left(x_{\max} \frac{k_B T}{\hbar}\right)^3}{e^{x_{\max}} - 1} \propto T^3$$

The relation that  $\hbar\omega_{\max} \propto T$  is known as Wien's displacement law. Combining these two proportionalities, we see that  $\text{Area} \sim \omega_{\max} j_{\omega_{\max}} \propto T^4$  which is the Stefan-Boltzmann law. We can derive the Stefan-Boltzmann law in another way:

$$\begin{aligned} U &= \int_0^\infty d\omega D_\omega(\omega) \langle \varepsilon_\omega \rangle \\ &= \pi \left(\frac{L}{\pi c}\right)^3 \int_0^\infty d\omega \frac{\hbar\omega^3}{e^{\beta\hbar\omega} - 1} \end{aligned}$$

---

<sup>3</sup>The non-monotonicity of  $xe^x$  means that the inverse is not single-valued. The upper real branch is given  $W_0$  while the lower real branch is  $W_1$ . This is, of course, a complex valued function.



$$\begin{aligned}
&= \pi \left( \frac{L}{\pi c} \right)^3 \int_0^\infty \frac{dx}{\beta \hbar} \frac{\hbar (x/\beta \hbar)^3}{e^x - 1} \\
&= \pi \left( \frac{L}{\pi c} \right)^3 \left( \frac{k_B T}{\hbar} \right)^4 \hbar \int_0^\infty dx \frac{x^3}{e^x - 1}
\end{aligned}$$

Even without integrating, we see that  $U \propto T^4$  and  $U \propto V$ . To get the multiplicative factor, we need to evaluate the integral. It turns out to have a nice form:

$$\begin{aligned}
&= \pi \left( \frac{L}{\pi c} \right)^3 \left( \frac{k_B T}{\hbar} \right)^4 \hbar \frac{\pi^4}{15} \\
&= \frac{\pi^2}{15 \hbar^3 c^3} V (k_B T)^4
\end{aligned}$$

Thus,

$$\mathcal{J}_u = \frac{1}{4} c \frac{U}{V} = \sigma T^4 \quad \sigma = \frac{\pi^2 k_B^4}{60 \hbar^3 c^3} \quad (24.8)$$

where  $\sigma$  is the *Stefan-Boltzmann constant*.

### Boltzmann's Derivation

When we integrate the frequency-resolved flux density, we obtain the Stefan-Boltzmann law, which states that  $u \sim T^4$ . Stefan discovered this law empirically, but Boltzmann managed to derive this law purely thermodynamically, long before quantum mechanics. He made three key assumptions which allowed him to do so. First, he assumed that electromagnetic radiation is extensive—double the volume, double the contained radiation. Second, he obtained from Maxwell's laws that the equation of state could be given  $P = \frac{1}{3}u$  (which is derived from, for example, the maxwell stress-energy tensor). Finally, he assumed that the chemical potential of a photon is equally zero, as photons leave with no cost.<sup>a</sup>

We can then write the energy of the system as

$$U(T, V, \mu) = U(T, v) = Vu(T)$$

The fact that  $\mu$  is equally zero means that we can remove the functional dependence. Further, the extensivity in  $V$  allows us to write the energy as linear in  $V$ . Further, in extensive systems, the Euler relation holds:

$$U = TS - PV + \mu N$$

The final term is again, equally zero. From Maxwell's laws, the pressure of EM radiation is given as a third the energy density, so  $PV = \frac{1}{3}U$ . Then, we can write

$$S = \frac{1}{T} (U + PV) = \frac{4}{3} \frac{1}{T} U = \frac{4}{3} \frac{Vu(T)}{T}$$

Finally, we use the maxwell relation

$$\left( \frac{\partial P}{\partial T} \right)_V = \left( \frac{\partial S}{\partial V} \right)_T$$

Substituting each side of the maxwell relation, we see that

$$\frac{1}{3}u'(T) = \frac{4}{3} \frac{u(T)}{T} \implies u'(T) = 4 \frac{u(T)}{T}$$

From this differential equation, we can easily see that

$$u \propto T^4$$

---

<sup>a</sup>This last assumption we will see holds true because the Planck distribution is a Bose-Einstein distribution with chemical potential zero. In fact, all photons have chemical potential zero.

# Chapter 25

## Harmonic Solid

Consider a chain of beads with separation  $a$ , which are each coupled to their neighbours by a spring with constant  $k$ . Thus, we consider the resting (mean) position of the  $j^{\text{th}}$  bead to be  $R_j = ja$ . However, in general, the beads are displaced by a position  $x_j$  from their resting point. Thus,

$$r_j = ja + x_j$$

The energy of such a system would be given

$$E = \sum_{j=1}^N \frac{1}{2} m \dot{x}_j^2 + \frac{1}{2} k (x_j - x_{j+1})^2$$

There are two ways to determine  $x_0$  and  $x_{N+1}$ : fixed boundary conditions or periodic boundary conditions. We will ignore the former and use the latter:  $x_{N+1} = x_1$ . Periodic boundary conditions are equivalent to putting the system on a torus. The benefit of periodic boundary conditions is that you have translational symmetry with no true boundary.

Luckily, even though the hamiltonian couples neighbouring particles, it is a quadratic form—we can rewrite in terms of a basis where the hamiltonian is diagonal, allowing us to decouple the hamiltonian. Recall that quadratic forms are necessarily hermitian, and thus can be diagonalized wrt an eigenbasis.

It turns out the eigenbasis for this hamiltonian is given by plane waves. We can write this basis as a discrete fourier transform

$$\tilde{x}_k = \frac{1}{\sqrt{N}} \sum_{j=1}^N x_j e^{-ikR_j} = \frac{1}{\sqrt{N}} \sum_{j=1}^N x_j e^{-ikaj}$$

where  $j$  is an integer, and the wavevector  $k = \frac{2\pi n}{L}$ . The inverse transformation is naturally another fourier transform

$$x_j = \frac{1}{\sqrt{N}} \sum_k \tilde{x}_k e^{+ikaj}$$

We can confirm this is the inverse transformation by plugging in the former to the latter:

$$\frac{1}{\sqrt{N}} \sum_k \frac{1}{\sqrt{N}} \sum_{j'} x_{j'} e^{ikaj'} e^{ikaj} = \sum_{j'} x_{j'} \frac{1}{N} \sum_k e^{ika(j-j')}$$

recognizing  $k = \frac{2\pi}{L}n = \frac{2\pi}{Na}n$ , we can rewrite

$$= \sum_{j'} x_{j'} \frac{1}{N} \sum_n e^{-\frac{2\pi}{N}n(j-j')}$$

If  $j \neq j'$ , we obtain a sum of the  $N^{\text{th}}$  roots of unity, which is equally zero. Thus, we have the exponential as a Kronecker delta,  $\frac{1}{N} \sum_k e^{ika(j-j')} = \delta_{j,j'}$ . Thus,

$$\begin{aligned} &= \sum_{j'} x_j \delta_{j,j'} \\ &= x_j \end{aligned}$$

Rewriting the Hamiltonian in Fourier space,

$$\begin{aligned} T &= \frac{1}{2}m \sum_j \dot{x}_j^2 \\ &= \frac{1}{2}m \sum_j \left( \frac{1}{\sqrt{N}} \sum_k \dot{\tilde{x}}_k e^{ikR_j} \right) \left( \frac{1}{\sqrt{N}} \sum_{k'} \dot{\tilde{x}}_{k'} e^{ik'R_j} \right) \\ &= \frac{1}{2}m \sum_{k,k'} \dot{\tilde{x}}_k \dot{\tilde{x}}_{k'} \frac{1}{N} \sum_j e^{i(k+k')R_j} \end{aligned}$$

Which we can rewrite again as a delta function, considering when the argument of the exponential is zero:

$$\begin{aligned} &= \frac{1}{2}m \sum_{k,k'} \dot{\tilde{x}}_k \dot{\tilde{x}}_{k'} \delta_{k,-k'} \\ &= \frac{1}{2}m \sum_k \dot{\tilde{x}}_k \dot{\tilde{x}}_{-k} \end{aligned}$$

Recalling the definition of  $\tilde{x}_k$ , we see that  $k \rightarrow -k$  is the same as  $\tilde{x}_{-k} = \tilde{x}_k^*$

$$= \frac{1}{2}m \sum_k |\dot{\tilde{x}}_k|^2$$

Similarly, we can rewrite the potential energy

$$\begin{aligned} U &= \frac{1}{2}K \sum_j (x_j - x_{j+1})^2 \\ &= \frac{1}{2}K \sum_j \left[ \frac{1}{\sqrt{N}} \sum_k \tilde{x}_k (e^{ikR_j} - e^{ikR_{j+1}}) \right] \left[ \frac{1}{\sqrt{N}} \sum_{k'} \tilde{x}_{k'} (e^{ik'R_j} - e^{ik'R_{j+1}}) \right] \\ &= \frac{1}{2}K \sum_j \left[ \frac{1}{\sqrt{N}} \sum_k \tilde{x}_k e^{ikaj} (1 - e^{ika}) \right] \left[ \frac{1}{\sqrt{N}} \sum_{k'} \tilde{x}_{k'} e^{ik'aj} (1 - e^{ik'a}) \right] \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2}K \sum_{k,k'} \tilde{x}_k (1 - e^{ika}) \tilde{x}_{k'} (1 - e^{ik'a}) \underbrace{\frac{1}{N} \sum_j e^{i(k+k')R_j}}_{=\delta_{k,-k'}} \\
&= \frac{1}{2}K \sum_k \tilde{x}_k \tilde{x}_{-k} (1 - e^{ika}) (1 - e^{-ika}) \\
&= \frac{1}{2}K \sum_k \tilde{x}_k \tilde{x}_k^* (2 - 2 \cos ka) \\
&= \sum_k \frac{1}{2}K(k) |\tilde{x}_k|^2
\end{aligned}$$

where we define

$$K(k) \equiv 4K \sin^2 \frac{ka}{2}$$

Thus, the potential is also diagonal in reciprocal space. Thus, we see that in reciprocal space, the modes are uncoupled. However, we see that the spring constant is dependent on the mode of the vibration in reciprocal space.

### Why reciprocal space?

Consider the original form of our potential:

$$U = \sum_j \frac{Ka^2}{2} \frac{(x_j - x_{j+1})^2}{a^2} \sim \frac{Ka^2}{2} \left( \frac{\partial u}{\partial x} \right)^2$$

We recognize the finite difference as a derivative. We can then consider it in terms of the  $\hat{p} = \frac{i}{\hbar} \frac{d}{dx}$ . The momentum eigenfunctions are plane waves  $\psi_p = e^{\pm ikx}$  which explains how the diagonal basis can be found in fourier space.

We can then find the dispersion relation of our harmonic solid as

$$\omega(k) = \sqrt{\frac{K(k)}{m}} = \sqrt{\frac{K}{m}} 2 \left| \sin \frac{ka}{2} \right| \equiv 2\tilde{\omega} \left| \sin \frac{ka}{2} \right|$$

Note we need only know the region  $k \in [-\frac{\pi}{a}, \frac{\pi}{a}]$ , which is known as the *First Brillouin Zone*. The points where we have valid wavevectors are spaced by  $2\pi/L$ . Note that the modes at the boundaries of the first Brillouin zone are the identical, as we have periodic boundary conditions. The first Brillouin zone is the only necessary domain we need to know the dispersion relation, as higher wavevectors are made redundant by the Sampling theorem and the Nyquist frequency.

For small  $k$ , the dispersion relation approaches  $\omega(k \ll 1) = \tilde{\omega}a |k|$ .

Note that there are two types of velocity. The first is called *phase velocity*, and determines how the phase of a wave moves. It is defined

$$v_p = \frac{\omega(k)}{k}$$

The other is the group velocity, and is the velocity of the envelope of a wavepacket

$$v_g = \frac{\partial \omega(k)}{\partial k} = c_s$$

where  $c_s$  is the speed of sound in the material. Notice that as  $k \rightarrow 0$ , we have  $v_g = v_p = \tilde{\omega}a$ .

## 25.1 Energy

### Classical Energy

The classical energy is trivial to determine, as there are  $N$  quadratic degrees of freedom for both position and momentum; thus, we should expect

$$U = Nk_B T$$

Thus, we will have the specific heat

$$c_v = \frac{1}{N} \left( \frac{\partial U}{\partial T} \right)_{V,N} = k_B \quad (25.1)$$

This last expression is known as the *Law of Dulong-Petit*. In a solid, there are 2 more degrees of freedom each for momentum and position, so in an ideal solid

$$c_v = 3k_B$$

which was what was measured at standard temperatures. However, when  $T \rightarrow 0$ , the measurements started diverging—the effects of quantum mechanics began to dominate.

### Quantum Energy

Because we were able to decouple the Hamiltonian in  $k$ -space, we can factorize the hamiltonian. Thus,

$$Z_{QM} = \prod_{j=1}^N Z_{Qm,k_j} = \prod_{j=1}^N \frac{1}{2 \sinh \frac{\beta \hbar \omega(k_j)}{2}} \quad (25.2)$$

The quantum free energy is then

$$F_{QM} = -k_B T \log Z_{QM} = k_B T \sum_{j=1}^N \log \left( 2 \sinh \frac{\beta \hbar \omega(k_j)}{2} \right) \quad (25.3)$$

the energy is

$$U_{QM} = \frac{\partial \beta F}{\partial \beta} = \sum_k \left[ \frac{\hbar \omega(k)}{2} + \frac{\hbar \omega(k)}{e^{\beta \hbar \omega(k)} - 1} \right] \quad (25.4)$$

Unfortunately, if we extend this to 3D, we would need to consider crystal symmetries, extra degrees of freedom, etc., and the sum becomes unreasonable to evaluate analytically

### 25.1.1 Debye Approximation

Debye developed an approximation to evaluate these sums. Restricting to low temperatures, we need only consider low  $\omega$ , so

$$\hbar\omega(\mathbf{k}) = \hbar\omega ak = \hbar c_s \frac{2\pi}{L} n$$

where

$$n = \|\mathbf{n}\| = \|\langle n_x, n_y, n_z \rangle\|$$

and

$$n_i = \frac{N}{2} + m \quad m \in [1, N] \subseteq \mathbb{N}$$

Further, we assume the Brillouin zone is spherical region that contains the  $3N$  modes we need for the waves in 3D. Further, note that the boundary of the Brillouin zone are at high energies/high temperatures, and thus their impact will not be very significant. Considering a continuum approximation of the states, we have

$$3N = 3 \int_0^{n_D/2} dn 4\pi n^2 \implies n_d = \left( \frac{6N}{\pi} \right)^{1/3}$$

The division by 2 is in analogy to the max value of  $n_i$  being  $N/2$ ; it is a sphere centred at the origin. Further, the 3 counts the independent modes a particle can vibrate; each wave has 2 transverse and one longitudinal mode.

The Debye energy is then

$$\begin{aligned} U_D &= 3 \int_0^{n_D/2} dn 4\pi n^2 \frac{\hbar\omega(n)}{e^{\beta\hbar\omega(n)} - 1} \\ &= 12\pi \int_0^{n_D/2} 4\pi n^2 \frac{\hbar c_s \frac{2\pi}{L} n}{e^{\beta\hbar c_s \frac{2\pi}{L} n} - 1} \quad x \equiv \beta\hbar c_s \frac{2\pi}{L} n \\ &= 12\pi \int_0^{\beta\hbar c_s \pi n_D/L} \frac{x^2 dx}{(\beta\hbar c_s \frac{2\pi}{L})^3} \frac{x/\beta}{e^x - 1} \\ &= \frac{12\pi}{\beta} \left( \frac{L}{\beta\hbar c_s 2\pi} \right)^3 \int_0^{\beta\hbar c_s \pi n_D/L} dx \frac{x^3}{e^x - 1} \\ &= \frac{3\pi}{2} \frac{L^3 (k_B T)^4}{(\pi\hbar c_s)^3} \int_0^{\Theta_D/T} dx \frac{x^3}{e^x - 1} \quad \Theta_d \equiv \frac{\hbar c_s \pi n_D}{L k_B} \end{aligned}$$

Unfortunately, this is an ugly finite integral, and is  $T$  dependent from the boundary. We can however, inspect two limits. If<sup>1</sup> we send  $T \rightarrow \infty$ , we can taylor expand our integrand to obtain

$$\frac{x^3}{e^x - 1} \rightarrow \frac{x^3}{1 + x + O(x^2) - 1} = \frac{x^2}{1 + O(x)} = x^2[1 - O(x)] = x^2 - O(x^3)$$

---

<sup>1</sup>Note that this contradicts our initial assumption that we enforce low temperature so our dispersion relation is approximately linear, but the result is still valid.

where we expand the denominator as a geometric series. Thus,

$$U_D(T \rightarrow \infty) \approx \frac{3\pi}{2} \frac{L^3 (k_B T)^4}{(\pi \hbar c_s)^3} \int_0^{\Theta_D/T} dx x^2 = 3Nk_B T$$

which is in fact, the classical result! Further, we see there is no  $\hbar$  in this expression. If we take the low temperature limit, we see

$$U_D(T \rightarrow 0) \approx \frac{3\pi}{2} \frac{L^3 (k_B T)^4}{(\pi \hbar c_s)^3} \int_0^\infty \frac{x^3}{e^x - 1} \Rightarrow \frac{U_D(T \rightarrow 0)}{V} = u \approx \frac{3\pi^2}{30\hbar^3 c_s^3} (k_B T)^4$$

It is important to recognize the 3 in the numerator because it comes from the number of independent modes a wave could have. Further, we see that this is very similar to the blackbody radiation, albeit with a factor of 2 rather than 3; this is because light only has 2 transverse modes but no longitudinal modes. Similarly, we replace  $c_s \rightarrow c$  because light travels at  $c$  through a solid. Once again, the factor of  $\hbar$  indicates we cannot understand this result classically. If we consider the heat capacity at low temperature, we see

$$c_V = \frac{1}{N} \left( \frac{\partial U}{\partial T} \right)_V \propto T^3$$

This temperature dependence has been experimentally observed, and could not be explained classically. Debye's explanation was another early triumph for quantum mechanics. The analogy between blackbody radiation and the Debye approximation shows that both systems are fundamentally the same—they are an ideal gas of bosons.



# Chapter 26

## Ideal Quantum Gas

### 26.1 Fermions and Bosons

Consider a wavefunction of  $N$  particles.

$$\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$$

There is no obvious permutation symmetries. However, if we swap two particles, say  $\mathbf{r}_i, \mathbf{r}_j$ , the system is physically the same, so the modulus squared must be conserved and the only difference between these two wavefunctions must be a phase. However, by performing such a swap twice, we should return to the original phase; thus, the only two possible phase differences are 0 and  $\pi$ —the wavefunction must be either be totally symmetric or totally antisymmetric with respect to those indices.<sup>1</sup> We define the symmetric case to be bosons, while the anti-symmetric case to be fermions.<sup>2</sup> Note, that fermions have half-integer spin, while bosons have full-integer spin. The connection between the spin properties and the statistical properties is known as the *Spin-Statistics Theorem*. We can easily (anti)symmetrize an arbitrary function. Our symmetrization for bosons is given:

$$\psi_B(\vec{r}) = X_B \sum_{\pi \in S(N)} \psi(\pi \vec{r})$$

and our antisymmetrization for fermions is given

$$\psi_F(\vec{r}) = X_F \sum_{\pi \in S(N)} \sigma(\pi) \psi(\pi \vec{r})$$

where  $\pi$  is an element in the permutation group  $S(N)$ ,  $\sigma(\pi)$  is the parity of the permutation  $\pi$ , and  $X_{F,B}$  is a normalization constant. In the special case where we can separate a wavefunction into a product of single particle wave functions,

$$\psi(\vec{r}) = \prod_i \psi_{\alpha_i}(r_i)$$

---

<sup>1</sup>Apparently *anyons* can exist, where the phase difference can take *any* value, but these exist solely as quasi-particles in 2D systems

<sup>2</sup>The *Pauli Exclusion Principle* arises trivially from the antisymmetry of the fermion wavefunction.

then, our (anti)symmetrization becomes

$$\psi_{F,B} = X_{F,B} \sum_{\pi} \{\sigma(\pi), 1\} \psi_{\alpha_i}(r_{\pi(i)})$$

In the Fermion case, this is known as the *Slater Determinant*

### 26.1.1 Partition Function

Recall from our discussion of the multi-particle partition function that if  $H = \sum_j H_j$  where all  $H_j$  are identical, we can write

$$H\psi_{F,B} = \sum_j \varepsilon_{\alpha_j} \psi_{F,B}$$

so

$$Z = \sum_{\{\alpha\}} e^{-\beta \sum_j \varepsilon_{\alpha_j}} = \sum_{\{\alpha\}} \prod_{j=1}^{\infty} e^{-\beta E_{\alpha_j}}$$

which is the sum over *all* quantum states. However, not all possible states are *valid* states that satisfy the necessary (anti)symmetry properties of Fermions and Bosons—this means that the partition function *does not factorize*. Instead we write

$$Z = \sum'_{\{\alpha\}} e^{-\beta \sum_j \varepsilon_{\alpha_j}} = \sum'_{\{\alpha\}} \prod_{j=1}^{\infty} e^{-\beta E_{\alpha_j}}$$

The prime indicates that this sum is subject to the (anti)symmetrization condition; the partition function does not factorize because the sum of one state depends on the other states, *even in the absence of interactions*.

It is possible, however, through a change in notation, to make this sum possible to do. Consider an arbitrary energy spectrum with arbitrary degeneracy with states denumerated  $n$ . There are two ways to characterize a configuration of particles. For example, we can list the state in which a list of particles is occupying, such as  $\{\alpha_j\} = \{1, 2, 2, 5, 7, 8\}$  stating that particle 1 is in state 1, particles 2,3 are in state 2, and so forth.<sup>3</sup> We can also describe a state by describing the occupancy of the states. For the same state as the previous example, we can write  $\{n_{\alpha}\} = \{1, 2, 0, 0, 1, 0, 1, 1, 0, \dots\}$ . We call  $n_{\alpha}$  the occupancy number of state  $\alpha$ . The advantage of this occupancy number representation is that we are no longer constrained by the indistinguishability of particles, we can instead write the partition function as

$$Z_N = \sum_{\{n_{\alpha}\}}^{\sum_{\alpha} n_{\alpha} = N} \prod_{\alpha} e^{-\beta \varepsilon_{\alpha} n_{\alpha}} \quad (26.1)$$

In this manner, the partition function neatly factorizes, *even for fermions*, where we restrict  $n_{\alpha} \in \{1, 0\}$ . However, we still have the constraint where  $\sum_{\alpha} n_{\alpha} = N$  for all  $\{n_{\alpha}\}$ . However, there is luckily an incredibly elegant solution for this constraint:  $N$  need not be fixed!

---

<sup>3</sup>Note this does not account for the indistinguishability of particles

Indeed we do have an ensemble where we need not fix  $N$ —the Grand Canonical Ensemble. Thus, we exchange fixed  $N$  for fixed  $\mu$ . Recall that the grand canonical partition function is the discrete Laplace transform of the canonical partition function:

$$\mathcal{Z} = \sum_{N=0}^{\infty} Z_N e^{\beta\mu N}$$

Thus, for our system, we can write the partition function as

$$\mathcal{Z} = \sum_{N=0}^{\infty} \sum_{\{n_\alpha\}}^{\sum_\alpha n_\alpha = N} \prod_\alpha e^{-\beta(\varepsilon_\alpha - \mu)n_\alpha}$$

We see after a little thought the Laplace transform and the constraint cancel each other out, leaving us with the sum over all possible configurations:

$$\mathcal{Z} = \sum_{\{n_\alpha\}} \prod_\alpha \exp[-\beta(\varepsilon_\alpha - \mu)n_\alpha] \quad (26.2)$$

expanding out the sums and products,

$$= \sum_{n_1} \sum_{n_2} \dots e^{-\beta(\varepsilon_1 - \mu)n_1} e^{\beta(\varepsilon_2 - \mu)n_2} \dots$$

Thus, we see that we can factorize the partition function as desired:

$$\begin{aligned} &= \sum_{n_1} e^{-\beta(\varepsilon_1 - \mu)n_1} \sum_{n_2} \dots \\ &= \prod_\alpha \sum_{n_\alpha} e^{-\beta(\varepsilon_\alpha - \mu)n_\alpha} \end{aligned} \quad (26.3)$$

We can then write our Grand potential as

$$-\beta\Omega = \ln \mathcal{Z} = \sum_\alpha \ln \sum_{n_\alpha} e^{-\beta(\varepsilon_\alpha - \mu)n_\alpha}$$

Now, it makes a difference between fermions and bosons, as for fermions, we have  $n_\alpha \in \{0, 1\}$ , while for bosons we have  $n_\alpha \in \mathbb{N} \cup \{0\}$ . The innermost sum for the fermion case yields the simple result

$$\sum_{n_\alpha=0}^1 = 1 + e^{-\beta(\varepsilon_\alpha - \mu)}$$

while for the boson case, we see that we can rewrite the sum as a geometric series, so

$$\sum_{n_\alpha=0}^{\infty} = \frac{1}{1 - e^{-\beta(\varepsilon_\alpha - \mu)}}$$

Note however, the geometric series only converges if  $\mu < \min_\alpha \{\varepsilon_\alpha\}$ , or the chemical potential must be lower than the ground state. Thus, we have the grand potential as

$$\beta\Omega_\pm = \mp \sum_\alpha \ln [1 \pm e^{-\beta(\varepsilon_\alpha - \mu)}] \quad (26.4)$$

where  $\Omega_+$  corresponds to Fermions, and  $\Omega_-$  corresponds to bosons.

## 26.2 Distribution Functions

Now that we have the partition function, we are interested in the average particles in a given state. We have

$$\begin{aligned}
 \langle n_\alpha \rangle &= \frac{1}{\mathcal{Z}} \left( \prod_{\alpha'} \sum_{n_{\alpha'}} n_{\alpha'} e^{-\beta(\varepsilon_{\alpha'} - \mu)n_{\alpha'}} \right) \\
 &= \frac{1}{\mathcal{Z}} \left[ \prod_{\alpha'} \sum_{n_{\alpha'}} \left( -\frac{1}{\beta} \frac{\partial}{\partial \varepsilon_{\alpha'}} \right) e^{-\beta(\varepsilon_{\alpha'} - \mu)n_{\alpha'}} \right] \\
 &= \frac{1}{\mathcal{Z}} \left( -\frac{1}{\beta} \frac{\partial}{\partial \varepsilon_{\alpha}} \right) \left[ \prod_{\alpha'} \sum_{n_{\alpha'}} e^{-\beta(\varepsilon_{\alpha'} - \mu)n_{\alpha'}} \right] \\
 &= \frac{1}{\mathcal{Z}} \left( -\frac{1}{\beta} \frac{\partial}{\partial \varepsilon_{\alpha}} \right) \mathcal{Z} \\
 &= \frac{\partial}{\partial \varepsilon_{\alpha}} (-k_B T \ln \mathcal{Z}) \\
 &= \frac{\partial \Omega}{\partial \varepsilon_{\alpha}} \\
 &= \frac{\partial}{\partial \varepsilon_{\alpha}} \left[ \mp k_B T \sum_{\alpha'} \ln (1 \pm e^{-\beta(\varepsilon_{\alpha'} - \mu)n_{\alpha'}}) \right] \\
 &= \pm k_B T \frac{\pm e^{-\beta(\varepsilon_{\alpha} - \mu)} (-\beta)}{1 \pm e^{-\beta(\varepsilon_{\alpha} - \mu)}}
 \end{aligned}$$

Thus, we obtain the the Fermi and Bose distributions:

$$\langle n_\alpha \rangle = \frac{1}{e^{\beta(\varepsilon_{\alpha} - \mu)} \pm 1} \quad (26.5)$$

We will often write using

$$f_{\pm}(\varepsilon) = \frac{1}{e^{\beta\varepsilon} \pm 1} \quad (26.6)$$

so our distributions become

$$\langle n_\alpha \rangle = f(\varepsilon_{\alpha} - \mu)$$

Plotting our functions, we see that the fermi distribution looks like a logistic curve, starting at 1, then crossing  $f_+(0) = \frac{1}{2}$  to become 0 at infinity. On the other hand, the bose distribution diverges at zero, which is where the chemical potential no longer satisfies the criterion  $\mu < \min_{\alpha} \{\varepsilon_{\alpha}\}$ . We see that at sufficiently high energy, the states are too high to be occupied. However, as the energy decreases wrt the temperature, the state approaches maximum occupancy, which for fermions is 1, while for bosons, there is no maximum.

## 26.3 Density of States

As sums are rather annoying to deal with, we will reformulate everything in terms of integrals over densities of state. Consider the differential DoS

$$D_\varepsilon = \sum_\alpha \delta(\varepsilon - \varepsilon_\alpha) \quad (26.7)$$

Note that after integration, this becomes *identical* to our discrete sum (we will consider continuous density of states later). Then, our cumulative density of states can be written

$$W(\varepsilon) = \int_{-\infty}^{\varepsilon} d\varepsilon' D_\varepsilon(\varepsilon') \quad (26.8)$$

Thus, we can, for instance, write the grand potential as

$$\Omega_\pm = \mp k_B T \int d\varepsilon D_\varepsilon(\varepsilon) \ln [1 \pm e^{-\beta(\varepsilon-\mu)}]$$

which is *equivalent* to our discrete sum. However, as an integral, we have far more techniques of manipulation—for instance, integration by parts.

$$\Omega_\pm = \mp k_B T \left[ W(\varepsilon) \ln [1 \pm e^{-\beta(\varepsilon-\mu)}] \Big|_0^\infty - \int d\varepsilon W(\varepsilon) f_\pm(\varepsilon - \mu) \right]$$

so,

$$\Omega_\pm = - \int d\varepsilon W(\varepsilon) f_\pm(\varepsilon - \mu) \quad (26.9)$$

Similarly, we can rewrite the average occupation

$$\begin{aligned} \langle N \rangle &= \sum_\alpha \langle n_\alpha \rangle \\ &= \int d\varepsilon D(\varepsilon) f_\pm(\varepsilon - \mu) = - \frac{\partial \Omega}{\partial \mu} \end{aligned}$$

where the last equality comes from differentiating our integral equation for the grand canonical potential then integrating by parts (or by basic thermodynamics). More explicitly, we can write

$$N(\mu) = \int d\varepsilon \frac{D(\varepsilon)}{e^{\beta(\varepsilon-\mu)} - 1} \quad (26.10)$$

This equation is frequently used to find  $\mu(N)$  if we wish to use  $N$  rather than  $\mu$ .

We can also find the average energy as

$$E = \sum_\alpha \varepsilon_\alpha \langle n_\alpha \rangle$$

so

$$E = \int d\varepsilon D(\varepsilon) \varepsilon f_\pm(\varepsilon - \mu) \quad (26.11)$$

We can rewrite this expression in terms of  $\Omega$  by using the following identity

$$\frac{\partial}{\partial \beta} \ln [1 \pm e^{-\beta(\varepsilon-\mu)}] = \frac{\mp e^{-\beta(\varepsilon-\mu)}}{1 \pm e^{-\beta(\varepsilon-\mu)}} (\varepsilon - \mu) = \mp (\varepsilon - \mu) f_{\pm}(\varepsilon - \mu) \quad (26.12)$$

Thus, we can rewrite

$$\begin{aligned} E &= \int d\varepsilon D(\varepsilon) \left( \mp \frac{\partial}{\partial \beta} \ln [1 \pm e^{-\beta(\varepsilon-\mu)}] + \mu f_{\pm}(\varepsilon - \mu) \right) \\ &= \left( \frac{\partial}{\partial \beta} \mp \int d\varepsilon D(\varepsilon) \ln [1 \pm e^{-\beta(\varepsilon-\mu)}] \right) + \mu \int d\varepsilon D(\varepsilon) f_{\pm}(\varepsilon - \mu) \\ &= \frac{\partial}{\partial \beta} (\beta \Omega) - \mu \frac{\partial \Omega}{\partial \mu} \end{aligned} \quad (26.13)$$

We recognize the first term as  $E - \mu N$  and the second term as  $+\mu N$ , as follows from thermodynamics and our previous derivation respectively. Because we have the grand potential  $\Omega(T, V, \mu)$ , as a result of our manipulation of the original sum, it is useful to have a way to write the other thermodynamic variables in terms of the one we have based all of our discussion on.

## 26.4 Model Systems

In order to analyze model systems, we need to determine the density of states as follows from the spectrum  $\{\varepsilon_{\alpha}\}$  of the Hamiltonian for those states.<sup>4</sup> Consider the spectrum of a particle in a  $L^d$  box. Of course, we have the energy as

$$\varepsilon = \frac{P^2}{2m} = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2}{2m} \left( \frac{\pi}{L} \mathbf{n} \right)^2 \quad N \in \mathbb{N}_0^d \setminus \{\vec{0}\}$$

To simplify the calculation, we will consider a continuous density of space—the density of  $\mathbf{n}$  in  $n$ -space is unity. By the transformation theorem, we can calculate the density of states in energy space:

$$\begin{aligned} D(\varepsilon) &= \int_{\mathbb{N}_0^d \setminus \{\vec{0}\}} d^d n \, 1 \delta \left( \varepsilon - \frac{\hbar^2}{2mL^2} n^2 \right) \\ &= \frac{1}{2^d} \int_{\mathbb{R}^d} dd n \, \delta \left( \varepsilon - \frac{\hbar^2 \pi^2}{2mL^2} n^2 \right) & dy &= \frac{\pi \hbar}{\sqrt{2mL^2}} dn \\ &= \frac{1}{2^d} \left( \frac{\sqrt{2mL^2}}{\hbar \pi} \right)^d \int_{\mathbb{R}^d} d^d y \, \delta(\varepsilon - y^2) \\ &= V \left( \frac{\sqrt{2m}}{2\pi \hbar} \right)^d \int_0^\infty dy \, A_d y^{d-1} \delta(\varepsilon - y^2) \end{aligned}$$

---

<sup>4</sup>Recall that this is not a general discussion of quantum statistical mechanics; we have assumed that our hamiltonian factorizes into a sum of individual, uncoupled particles with identical spectra.

Where, of course,  $A_d$  is the surface area of a  $d$ -dimensional sphere. Substituting  $x = y^2$ , we can rewrite as

$$\begin{aligned} &= \frac{1}{2^d} \left( \frac{\sqrt{2m}}{\pi \hbar} \right)^d V \int_0^\infty \frac{dx}{2\sqrt{x}} A_d x^{(d-1)/2} \delta(\varepsilon - x) \\ &= \frac{1}{2} \frac{(2m)^{d/2}}{h^d} V A_d \varepsilon^{\frac{d}{2}-1} \end{aligned}$$

Recall, of course, that we have  $A_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}$

$$= \left( \frac{\sqrt{2\pi m}}{h} \right)^d \frac{V}{\Gamma(d/2)} \varepsilon^{\frac{d}{2}-1}$$

If we have spin angular momentum, we multiply this by a factor of the spin degeneracy,  $(2s + 1)$ .

### 26.4.1 Equation of Clapeyron

There is an older result that can be found to be the direct consequence of  $D_E \sim E^{\frac{d}{2}-1}$ . We see that we can write

$$D(E) = \frac{2}{d} \frac{\partial}{\partial E} [ED(E)]$$

This can be verified by expanding the RHS. Recall that for extensive systems we have  $PV = -\Omega$ . We have an expression for  $\Omega$ , which allows us to rewrite

$$\begin{aligned} PV &= \pm k_B T \int_0^\infty d\varepsilon D(\varepsilon) \ln(1 \pm e^{-\beta(\varepsilon-\mu)}) \\ &= \pm k_B T \frac{2}{d} \int_0^\infty d\varepsilon \left( \frac{\partial}{\partial \varepsilon} [\varepsilon D(\varepsilon)] \right) \ln(1 \pm e^{-\beta(\varepsilon-\mu)}) \end{aligned}$$

Integrating by parts (and noticing the boundary terms vanish),

$$\begin{aligned} &= \mp k_B T \frac{2}{d} \int_0^\infty d\varepsilon \varepsilon D(\varepsilon) \frac{\mp \beta e^{-\beta(\varepsilon-\mu)}}{1 \pm e^{-\beta(\varepsilon-\mu)}} \\ &= \frac{2}{d} \int_0^\infty d\varepsilon \varepsilon D(\varepsilon) f_\pm(\varepsilon - \mu) \\ &= \frac{2}{d} E \end{aligned}$$

Thus, we obtain the equation of Clapeyron:

$$E = \frac{d}{2} PV$$

It is amazing that considering quantum contributions, we obtain the same relation as the classical result even though the ideal gas law,  $PV = Nk_B T$ , and equipartition theorem,  $E = \frac{d}{2} Nk_B T$  do *not* hold in quantum mechanics!

### 26.4.2 Grand Potential

Recall our expression for the grand potential,

$$\Omega = \mp k_B T \int_0^\infty d\varepsilon \varepsilon D(\varepsilon) \ln(1 \pm e^{-\beta(\varepsilon-1)})$$

where we define the ground state  $\varepsilon_0 = 0$ . It is often useful to define (in both classical and quantum statistical mechanics) a quantity

$$z = e^{\beta\mu} \quad (26.14)$$

known as the *fugacity*. This allows us to evaluate

$$\begin{aligned} \Omega_\pm &= \mp k_B T \int_0^\infty d\varepsilon (2s+1) \left( \frac{\sqrt{2\pi m}}{h} \right)^d \frac{V}{\Gamma(\frac{d}{2})} \varepsilon^{\frac{d}{2}-1} \ln(1 \pm z e^{-\beta\varepsilon}) \\ &= \mp k_B T (2s+1) \left( \frac{\sqrt{2\pi m}}{h} \right)^d \frac{V}{\Gamma(\frac{d}{2})} \int_0^\infty d\varepsilon \varepsilon^{\frac{d}{2}-1} \ln(1 \pm z e^{-\beta\varepsilon}) \end{aligned}$$

Substituting  $t = \beta\varepsilon$ ,

$$\begin{aligned} &= \mp k_B T (2s+1) \left( \frac{\sqrt{2\pi m}}{h} \right)^d \frac{V}{\Gamma(\frac{d}{2})} \frac{1}{\beta^{d/2}} \int_0^\infty dt t^{\frac{d}{2}-1} \ln(1 \pm z e^{-t}) \\ &= \mp k_B T (2s+1) \left( \frac{\sqrt{2\pi m k_B T}}{h} \right)^d V \frac{1}{\Gamma(\frac{d}{2})} \int_0^\infty dt t^{\frac{d}{2}-1} \ln(1 \pm z e^{-t}) \end{aligned}$$

We recognize the integral with the factor of the Gamma as a polylog, and the term in parentheses corresponds to the thermal de Broglie wavelength. Thus,

$$\Omega_\pm = \pm k_B T (2s+1) \frac{V}{\lambda_{\text{th}}^d} L_{\frac{d}{2}+1}(\mp z) \quad (26.15)$$

From this, we can obtain two important relations. First,

$$\frac{PV}{k_B T} = -\beta\Omega = \mp \frac{V}{\lambda_{\text{th}}^d} (2s+1) L_{\frac{d}{2}+1}(\mp z) \quad (26.16)$$

Second, using the fact that

$$\begin{aligned} z \frac{\partial}{\partial z} &= z \frac{\partial \mu}{\partial z} \frac{\partial}{\partial \mu} = z \left( \frac{\partial \ln z}{\partial z} \frac{1}{\beta} \right) \frac{\partial}{\partial \mu} = \frac{1}{\beta} \frac{\partial}{\partial \mu} \\ z \frac{\partial}{\partial z} \beta &= \frac{\partial}{\partial \mu} \end{aligned}$$

we can make use of the ladder property of the polylog to obtain

$$N = -\frac{\partial \Omega}{\partial \mu} = z \frac{\partial}{\partial z} (-\beta\Omega) = \mp \frac{V}{\lambda_{\text{th}}^d} (2s+1) L_{\frac{d}{2}}(\mp z) \quad (26.17)$$



We can view these two equations as a parametric representation of the thermal equation of state, with  $z$  being the parameter. Dividing these two equations, we obtain

$$\frac{PV}{Nk_B T} = \frac{L_{\frac{d}{2}+1}(\mp z)}{L_{\frac{d}{2}}(\mp z)} \quad (26.18)$$

From the facts that

$$\begin{aligned} L_\nu(0) &= 0 & \frac{d}{dz} L_\nu &> 0 \\ L'_\nu(0) &= 1 & \frac{d}{d\nu} L_\nu &< 0 \end{aligned}$$

we see that for bosons we have

$$\frac{PV}{Nk_B T} \leq 1$$

while for fermions we will have

$$\frac{PV}{Nk_B T} \geq 1$$

Finally, for our classical ideal gas (boltzmannions) we have exactly<sup>5</sup>

$$\frac{PV}{Nk_B T} = 1$$

Thus, fermions exhibit a greater pressure than we would classically expect, while bosons exhibit a weaker pressure than we would classically expect. We can interpret this as Fermi (Bose) statistics leads to an addition repulsion (attraction) between particles. While these particles do not interact through any forces, the exchange statistics lead to what we can view as an interaction between particles.

---

<sup>5</sup>Note that this occurs on the polylogs at  $z = 0$ , so the classical limit is  $z \rightarrow 0$ .



# Chapter 27

## Bose-Einstein Condensation

### 27.1 Macroscopic Occupation

In a Bose gas with a non-interacting hamiltonian, as  $T \rightarrow 0$ , we expect that all the particles occupy the same ground state. However, this is not remarkable; it is not even what a Bose-Einstein condensate (BEC) is. What makes a BEC special is that for a given density  $n = N/V$ , there exists a temperature  $T_E > 0$  such that for  $T \leq T_E$ , the ground state is *macroscopically occupied*, or that the density of the ground state  $n_0 = \lim_{V \rightarrow \infty} \langle n_{\alpha_0} \rangle / V > 0$ . That is, a finite percent of the particles exist in the ground state, and the occupation in all other states go to zero.

Recall that we have

$$\langle n_{\alpha_0} \rangle = \frac{1}{e^{\beta(\varepsilon_0 - \mu)} - 1} = \frac{1}{z - 1} = \frac{z}{1 - z} \sim V = L^d \quad (27.1)$$

where we take the ground state to be  $\varepsilon_0 \rightarrow 0$  and  $z = e^{\beta\mu}$  is the fugacity. What follows from this is that we must have

$$z \sim \frac{\alpha L^d}{1 + \alpha L^d} \approx 1 - O(L^{-d}) = 1 - c/V \approx 0$$

or, the chemical potential must approach  $0 = \varepsilon_0$  from below, which is the convergence limit!.

If we consider the next excited state,  $\varepsilon_k \sim k^2 \equiv \delta k^2 / L^2$ , we see

$$\langle n_{\alpha_1} \rangle = \frac{1}{e^{\beta\varepsilon - \mu} - 1} = \frac{z}{e^{\beta\delta/L^2} - z} \approx \frac{1 - cL^{-d}}{\beta\delta/L^2 + cL^{-d}} = \frac{L^d - c}{\beta\delta L^{d-2} + c} \sim L^2$$

Taking  $\langle n_{\alpha_{n \geq 1}} \rangle / L^d \sim L^{-(d-2)}$ , we see that in 3D this trivially goes to zero. However, for 1 and 2D, this ratio becomes a constant.<sup>1</sup> As such, when we integrate/sum over all states, we see that the integral/sum diverges, which is a non-physical result. Thus, we cannot have

---

<sup>1</sup>For 1D, the equation given  $L^{-(d-2)}$  is not valid, as the denominator is instead dominated by  $c$ , so the relation instead becomes  $\langle n_{\alpha_{n \geq 1}} \rangle \sim L^{d-1}$ .

a finite fraction in the ground state, and correspondingly, we cannot have BECs in 1,2D. However, while we showed that there is no contradiction for BECs to exist in 3D, we have not shown that such macroscopic occupation *does* exist.

## 27.2 Condensation

Recall our expression for the integral density of states,

$$N = (2s + 1) \frac{V}{\lambda_{\text{th}}^d} L_{\frac{d}{2}}(z) \quad (27.2)$$

Recall from our discussion that  $z$  is bounded above by 1. Further, we showed that  $L_\nu(1) = \zeta(\nu)$ . Thus, because the polylog is monotonic, we see that we must have

$$N \leq (2s + 1) \frac{V}{\lambda_{\text{th}}^d} \zeta\left(\frac{d}{2}\right) \quad (27.3)$$

We know that for  $d = 1, 2$  the polylog diverges as  $z \uparrow 1$ , but *converges* for  $d = 3$ , as we have  $\zeta\left(\frac{3}{2}\right) \approx 2.612$ . Thus, in three dimensions, we expect to have

$$N \leq (2s + 1) \frac{V}{\lambda_{\text{th}}^3} \cdot 2.612 \dots$$

bounded above by a finite number (for a finite volume/temperature). For a bose system, we should not expect there to be any bounds, as an infinite number of bosons should, in principle, be able to fit in any state. There are no repulsions, and the particles have no volume, and so there should not be any upper bound.

This contradiction arises because we used a continuous model for the density of states to derive  $N$  rather than the true discrete energy spectrum. The “smearing out” of the DoS is generally a valid approximation, but it fails for the ground state. Recall, that we had in 3D that  $D(\varepsilon) \sim \sqrt{\varepsilon}$ , which *destroys the macroscopically occupied state* at  $\varepsilon = 0$ . In other words, if we miss the ground state particles, we miss a finite fraction of all particles.

Thus, our integrated density of states measures just the number of particles in *excited* states:

$$N_{\text{excited}} = (2s + 1) \frac{V}{\lambda_{\text{th}}^d} L_{\frac{d}{2}}(z) \leq (2s + 1) \frac{V}{\lambda_{\text{th}}^d} \zeta\left(\frac{d}{2}\right) \quad (27.4)$$

When the number of particles in our system then exceeds this number, *the rest must go to the ground state*. Thus, we observe macroscopic occupation of the ground state! This point marks the transition from a bose gas to a BEC.

We can thus calculate the ground state occupation,  $N_0$  as trivially

$$N_0 = N - (2s + 1) \frac{V}{\lambda_{\text{th}}^d} \zeta\left(\frac{d}{2}\right) \quad (27.5)$$

Trivially, we see that the macroscopic occupation is then given

$$\frac{N_0}{N} = 1 - (2s + 1)\zeta\left(\frac{d}{2}\right) \frac{V}{N\lambda_{th}^d} \quad (27.6)$$

In 3D, and reinserting the definition of  $\lambda_{th} = \frac{h}{\sqrt{2\pi mk_B T}}$ , we see that

$$\frac{N_0}{N} = 1 - \left(\frac{T}{T_E}\right)^{3/2}$$

where the einstein temperature is given

$$T_E = \frac{h^2}{2\pi mk_B} \left[ (2s + 1) \frac{V}{N} \zeta\left(\frac{3}{2}\right) \right]^{2/3} \quad (27.7)$$

For  $T < T_E$ , we have macroscopic ground state occupation. However, for  $T > T_E$ , we cannot have  $N_0/N < 0$ , so we fix  $N_0/N = 0$ , and we no longer have macroscopic occupancy of the ground state. There is trivially a discontinuity in the derivative of this fraction, so we must have a phase transition. At this critical temperature, we see that

$$(2s + 1)\zeta\left(\frac{3}{2}\right) \frac{1}{n\lambda_{T_E}^3} = 1$$

or

$$n\lambda_T^3 = \frac{1}{\zeta\left(\frac{3}{2}\right)(2s + 1)} \approx \frac{0.38}{2s + 1}$$

which is a very large density. In comparison, at STP, nitrogen has a thermal wavelength

$$\lambda = \frac{h}{\sqrt{2\pi mk_B T}} \approx 1.9 \times 10^{-10} \text{ m}$$

Thus, we would expect the particle density to be

$$n = \frac{N}{V} = \frac{P}{k_B T} \approx \frac{10^5 \text{ Pa}}{4.1e - 31} J = 2.4 \times 10^{25} \text{ m}^{-3}$$

so

$$n\lambda_{th}^3 = 1.6 \times 10^{-7} \ll 1$$

which is *tiny* compared to what we need for a BEC.

There are two ways we can achieve a BEC: we can either increase the density manually (which would make interactions important), or decrease the temperature to increase  $\lambda_{th}$ .

## 27.3 Properties of a BEC

Once a BEC forms, we have the fugacity becomes  $z \approx 1$  with incredible accuracy. Thus, we have

$$\frac{PV}{k_B T} = -\beta\Omega = (2s + 1) \frac{V}{\lambda_{th}^3} L_{\frac{5}{2}}(z) \approx (2s + 1) \frac{V}{\lambda_{th}^3} \zeta\left(\frac{5}{2}\right)$$

so

$$P = (2s + 1)\zeta\left(\frac{5}{2}\right) \cdot \frac{k_B T}{\lambda_{th}^3} \propto T^{5/2} \quad (27.8)$$

or, the pressure doesn't depend on the volume! Thus,

$$\left(\frac{\partial P}{\partial V}\right)_T = 0 \implies \kappa = -\frac{1}{V} \left(\frac{\partial V}{\partial P}\right) = \infty$$

so the system is *infinitely compressible*. We recognize another case with an infinitely compressible system as the phase transition between a liquid and a gas; at this transition, the system has constant pressure regardless of volume, as the gas which would have exerted a pressure condenses into a liquid. This analogy is why a BEC is called a *condensate*—as the volume decreases, we have more of the gas condensing into a BEC.

We can also calculate the heat capacity,

$$c_V = \frac{1}{N} \left(\frac{\partial U}{\partial T}\right)_V = \frac{3}{2} \frac{V}{N} \left(\frac{\partial P}{\partial T}\right)_V$$

so

$$c_V/k_B = \frac{15}{4} \frac{\zeta(\frac{5}{2})}{\zeta(\frac{3}{2})} \left(\frac{T}{T_E}\right)^{3/2} \quad (27.9)$$

where the constant prefactor is  $\approx 1.926$ . Above  $T_E$ , the heat capacity decays, so the Bose gas has highest heat capacity at  $T = T_E$  of  $c_V/k_B \approx 1.926$ .

Finally, at  $T = T_E$ , we can write

$$\begin{aligned} P &= (2s + 1)k_B T_E \frac{1}{\lambda_{th}^3} \zeta\left(\frac{5}{2}\right) \\ &= (2s + 1) \frac{h^2}{2\pi m} \frac{2\pi m k_B T_E}{h^2} \frac{1}{\lambda_{th}^3} \zeta\left(\frac{5}{2}\right) \\ &= (2s + 1) \frac{hbar^2}{2m} \frac{1}{\lambda_{th}^5} 4\pi \zeta\left(\frac{5}{2}\right) \\ &= (2s + 1) \frac{\hbar^2}{2m} \left[ (2s + 1) \zeta\left(\frac{3}{2}\right) \frac{N}{V} \right]^{5/3} 4\pi \zeta\left(\frac{5}{2}\right) \\ &= (2s + 1)^{8/3} \frac{\hbar^2}{2m} \left(\frac{N}{V}\right)^{5/3} 4\pi \zeta\left(\frac{5}{2}\right) \left[\zeta\left(\frac{3}{2}\right)\right]^{5/3} \end{aligned}$$

which is the transition line in the PV diagram of the BEC. Above this line, we have a normal gas, while below, we have the two-phase system. Plotting isotherms, we see in the gas phase that  $P \sim 1/V$ , which will intersect the transition line, then become a constant as  $V \rightarrow 0$ .

# Chapter 28

## Ideal Fermi Gases

Recall the Fermi-Dirac distribution function

$$f_+(\varepsilon - \mu) = \frac{1}{e^{-\beta(\varepsilon - \mu)} + 1}$$

For sufficiently large  $\varepsilon$ , we obtain the gibbs/boltzmann factor  $e^{-\beta\varepsilon}e^{+\beta\mu}$ , and similarly we obtain  $1 - e^{\beta(\varepsilon - \mu)}$ . At  $\varepsilon = \mu$ , we have  $f = 1/2$  and a slope of  $-\beta/4$ .

Our basic equations for the discussion of the fermi gas are

$$N = \sum_{\alpha} \langle n_{\alpha} \rangle = \sum_{\alpha} \frac{1}{e^{\beta(\varepsilon_{\alpha} - \mu)} + 1} = \int d\varepsilon \frac{D(\varepsilon)}{e^{\beta(\varepsilon - \mu)} + 1}$$

which allows us to find  $\mu(N)$ , and

$$U = \sum_{\alpha} \varepsilon_{\alpha} \langle n_{\alpha} \rangle = \sum_{\alpha} \frac{\varepsilon_{\alpha}}{e^{\beta(\varepsilon_{\alpha} - \mu)} + 1} = \int d\varepsilon \frac{\varepsilon D(\varepsilon)}{e^{\beta(\varepsilon - \mu)} + 1}$$

### Fermi Energy

We define the Fermi Energy  $\varepsilon_F$  to be the zero temperature limit of the chemical potential.

$$\varepsilon_F \equiv \lim_{T \rightarrow 0} \mu(T, N) \quad (28.1)$$

It is *not* the energy at which the Fermi distribution is 1/2, as that is the chemical potential. This is often a good approximation, however, as for small temperature, the corrections are correspondingly small.

We also have a rule-of-thumb regarding the Fermi energy—it lies halfway between the Highest Occupied state and the Lowest Unoccupied state (at, of course  $T = 0$ ). Thus, if we have  $N$  particles, the fermi energy is the energy between  $\varepsilon_N$  and  $\varepsilon_{N+1}$ .

Finally, we have a useful identity

$$\frac{1}{1+x} + \frac{1}{1+x^{-1}} = 1$$

which allows us to rewrite

$$\frac{1}{e^{\beta(\varepsilon-\mu)} + 1} = 1 - \frac{1}{1 + e^{-\beta(\varepsilon-\mu)}}$$

### Two-level system

Consider, for instance the two-level system with spectrum  $\{0, \varepsilon\}$ . Trivially, we can write

$$N = \frac{1}{e^{-\beta\mu} + 1} + \frac{1}{e^{\beta(\varepsilon-1)} + 1}$$

Plugging in our useful identity,

$$= 1 - \frac{1}{e^{+\beta\mu} + 1} + \frac{1}{e^{\beta(\varepsilon-\mu)} + 1}$$

Let us fix  $N = 1$ . Then, we see

$$\frac{1}{e^{\beta\mu} + 1} = \frac{1}{e^{\beta(\varepsilon-\mu)} + 1}$$

$$\mu = \varepsilon - \mu$$

$$\mu = \frac{1}{2}\varepsilon = \varepsilon_F$$

so we see that the chemical potential in the two-level system is always the same as the Fermi level, and *does not depend on temperature*. Further, we note that this agrees with our rule-of-thumb for the value of the Fermi energy.

## 28.1 Free Particle

Consider the spectrum of a free particle, which has the DoS of

$$D(\varepsilon) = (2s+1) \left( \frac{\sqrt{2\pi m}}{h} \right)^d \frac{V}{\Gamma(\frac{d}{2})} \varepsilon^{\frac{d}{2}-1}$$

Thus, at  $T = 0$ , we can write

$$N = \int_0^{\varepsilon_F} d\varepsilon D(\varepsilon) \tag{28.2a}$$

$$\begin{aligned} &= (2s+1) \left( \frac{\sqrt{2\pi m}}{h} \right)^d \frac{V}{\Gamma(\frac{d}{2})} \frac{2}{d} \varepsilon_F^{d/2} = (2s+1) \left( \frac{\sqrt{2\pi m \varepsilon_F}}{h} \right)^d \frac{V}{\frac{d}{2} \Gamma(\frac{d}{2})} \\ &\equiv \frac{2}{\Gamma(\frac{d}{2} + 1)} \frac{V}{\lambda_{\text{th},F}^d} \end{aligned} \tag{28.2b}$$



where we define

$$T_F = \frac{\varepsilon_F}{k_B} \quad (28.3)$$

in a characteristic Fermi thermal wavelength  $\lambda_{\text{th},F}$ .

In 3D, we have

$$N = \frac{V}{3\pi^2} \left( \frac{2m}{\hbar^2} \right)^{3/2} \varepsilon_F^{3/2}$$

so

$$\varepsilon_F = \frac{\hbar^2}{2m} (3\pi^2 n)^{2/3}$$

where  $n = N/V$ .

Similarly, we can obtain the energy (at  $T = 0$ !)

$$U = \int_0^{\varepsilon_F} d\varepsilon \varepsilon D(\varepsilon) \quad (28.4)$$

$$\begin{aligned} &= \int_0^{\varepsilon_F} d\varepsilon 2 \left( \frac{\sqrt{2\pi m}}{h} \right)^d \frac{V}{\gamma(\frac{d}{2})} \varepsilon^{d/2} \\ &= 2 \left( \frac{\sqrt{2\pi m}}{h} \right)^d \frac{V}{\Gamma(\frac{d}{2})(\frac{d}{2} + 1)} \varepsilon_F^{\frac{d}{2}+1} \end{aligned} \quad (28.5)$$

Dividing by  $N$ , we obtain (at  $T = 0$ )

$$u \equiv \frac{U}{N} = \frac{d}{d+2} \varepsilon_F \quad (28.6)$$

Using the Clapeyron equation, we can find the Fermi pressure at  $T = 0$

$$P = 2 \left( \frac{\sqrt{2\pi m}}{h} \right)^d \frac{\varepsilon_F^{\frac{d}{2}+1}}{\Gamma(\frac{d}{2})(1 + \frac{2}{d})} \quad (28.7)$$

Because  $\varepsilon_F \sim n^{2/d}$ , we have  $P \sim n^{1+\frac{2}{d}} \sim 1/V^{1+\frac{2}{d}}$ . More explicitly, at  $T = 0$  and  $d = 3$ , we obtain

$$P = 2 \left( \frac{\sqrt{2\pi m}}{h} \right)^3 \frac{[\frac{\hbar^2}{2m} (3\pi^2 n)^{2/3}]^{5/2}}{\frac{5}{3}\Gamma(\frac{3}{2})} = \frac{9}{10} (3\pi^2)^{2/3} \frac{\hbar^2}{2m} n^{5/3}$$

## Electrons in a metal

Consider electrons inside of a metal as an example of a Fermi gas<sup>1</sup>. In particular, we will assume that copper has valence of one. We can then calculate the number density (of free electrons) in copper as  $n \sim 8.5 \times 10^{28} \text{ m}^{-3}$ . The mass of the electrons is  $m_e = 9.11 \times 10^{-31} \text{ kg}$ .

---

<sup>1</sup>Why we can (considering charge, potential, etc) is entirely non-obvious; however, we will see later why we can do so.

We now have all of the information we need to calculate the zero-temperature fermi pressure as

$$P = 86 \text{ GPa} \approx 8.6 \times 10^5 \text{ bar}$$

which begs the question: how does copper not explode? We haven't even included the electrostatic repulsions of the other electrons, nor the fermi pressure of the nuclei! With regard to the latter, we see that the mass of the nuclei makes their fermi pressure negligible, while with regard to the former, it is actually the electrostatic attractions which holds the metal together.

If we consider the isothermal compressibility,

$$\kappa_T = \frac{1}{V} \left( \frac{\partial V}{\partial P} \right)_{T,N}$$

or rather, the isothermal bulk modulus (incompressibility),

$$K = -V \left( \frac{\partial P}{\partial V} \right)_{T,N}$$

we find

$$K = -V \left( -1 - \frac{2}{d} \right) \frac{P}{V} = \left( 1 + \frac{2}{d} \right) P$$

rewriting,

$$K = \left( 1 + \frac{2}{d} \right) \frac{PV}{N} \cdot \frac{N}{V} = \frac{d+2}{d} \frac{2}{d} \frac{u}{n}$$

inserting our expression for the energy density, we obtain

$$K = \frac{2}{d} \varepsilon_F n \quad (28.8)$$

Thus zero temperature bulk modulus of a free fermi gas is  $2/d$  times the fermi energy per particle volume<sup>2</sup>. For example, calculating the bulk modulus for the free electron gas in copper gives us 143 GPa; the actual bulk modulus is 123 GPa, so this is *very* close. This shows us the majority of the rigidity of materials comes not from electrostatic repulsions, but from the Pauli exclusion principle; squishing a material forces electrons into electrons to go into states they do not want to be.

## 28.2 Sommerfeld Expansion

We have focused on the case  $T = 0$ , but this is in general not useful. For small  $T$ , the thermodynamics is determined by the density of states near the Fermi energy. At  $T = 0$ ,

---

<sup>2</sup>Interestingly, the bulk modulus of *virtually anything* is given by the physically relevant energy scale divided by the physically relevant volume corresponding to that scale—up to a prefactor of order  $\sim 1$ . Considering bulk moduli often vary by orders of magnitude over a given range, this is a very useful approximation.

the fermi distribution is an exact step function; for small, finite  $T$ , this transition goes over a small region near the fermi energy, with width as a function of temperature.

We will consider an incredibly general problem. We wish to evaluate

$$I = \int d\varepsilon g(\varepsilon) f_+(\varepsilon - \mu)$$

for small  $T$ . The function  $g$  could be a density of states, a factor for an energy expectation, etc. In order to evaluate this integral, we employ the *Sommerfeld expansion*. We will make the following assumptions. First,

$$\lim_{\varepsilon \rightarrow -\infty} g = 0$$

or, the function is bounded below. Further, we wish for the function  $g(\varepsilon)$  to grow at most like a power law

$$g(\varepsilon) = \mathcal{O}(\varepsilon^n)$$

so after dividing by an exponential, the entire function vanishes and remains integrable. We must have  $g$  to be “sufficiently smooth” at the fermi energy  $\varepsilon = \varepsilon_F$  as we will be expanding around this point. Finally, we let

$$G(\varepsilon) \equiv \int_{-\infty}^{\varepsilon} d\varepsilon' g(\varepsilon') \iff G'(\varepsilon) = g(\varepsilon)$$

First, let  $\mu$  be fixed. Then,

$$I = \int_{-\infty}^{\infty} d\varepsilon G'(\varepsilon) \tilde{f}\left(\frac{\varepsilon - \mu}{k_B T}\right) \quad \tilde{f}(x) \equiv \frac{1}{e^x + 1}$$

Integrating by parts,

$$= \square - \int_{-\infty}^{\infty} d\varepsilon G(\varepsilon) \tilde{f}'\left(\frac{\varepsilon - \mu}{k_B T}\right) \frac{1}{k_B T}$$

We can throw out the boundary term because  $G$  vanishes at  $-\infty$  and grows slower than  $\tilde{f}$  vanishes at  $+\infty$ . Substituting  $x \equiv \frac{\varepsilon - \mu}{k_B T}$ ,

$$= - \int_{-\infty}^{\infty} dx G(\mu + k_B T x) \tilde{f}'(x)$$

We can now expand  $G$  for small  $T$ ,

$$\approx - \int_{-\infty}^{\infty} dx \tilde{f}'(x) \left[ G(\mu) + x k_B T G'(\mu) + \frac{1}{2} (x k_B T)^2 G''(\mu) + \mathcal{O}(T^3) \right]$$

Note that while we integrate over all  $x$ , the behaviour of  $\tilde{f}'(x) = -\frac{1}{4 \cosh^2(x/2)}$ , only values near  $x = 0$  are relevant; the cosh grows exponentially symmetrically about  $x = 0$ , so  $\tilde{f}'$  acts as a sort-of finite, wide delta function.

$$= G(\mu) \int_{-\infty}^{\infty} \tilde{f}'(x) dx - k_B T G'(\mu) \int_{-\infty}^{\infty} x \tilde{f}'(x) dx$$

$$-\frac{1}{2}(k_B T)^2 G''(\mu) \int_{-\infty}^{\infty} dx x^2 \tilde{f}'(x) - \mathcal{O}(T^3)$$

The first term can be calculated trivially, as the boundary terms of  $\tilde{f}(x)$ . The second requires more thought, but is a symmetric integral of an odd function, and thus vanishes; in fact, all even terms vanish. The third term is far more difficult, but using special functions, we can find it as  $-\pi^2/3$ . Thus, we can write

$$= G(\mu) + \frac{\pi^2}{6}(k_B T)^2 G''(\mu) + \mathcal{O}(T^4)$$

We can interpret this as resulting from a low-temperature expansion of the Fermi distribution at fixed  $\mu$ :

$$f_+(\varepsilon - \mu) = \Theta(\mu - \varepsilon) - \frac{\pi^2}{6}(k_B T)^2 \delta'(\varepsilon - \mu) + \mathcal{O}(T^4)$$

where  $\Theta$  is the *Heaviside function* or unit step function and  $\delta'$  is the derivative of the delta function. We can justify this by plugging our “expansion” into the integral, where we will obtain the Sommerfeld expansion.

### 28.2.1 Chemical Potential

Using the Sommerfeld potential, we can find  $\mu(T)$  for a given  $N$ . Let

$$W(\varepsilon) = \int_{-\infty}^{\varepsilon} d\varepsilon' D(\varepsilon')$$

be the integrated density of states. We can find  $N$  as

$$N = \int d\varepsilon D(\varepsilon) f_+(\varepsilon - \mu) = W(\mu) + \frac{\pi^2}{6}(k_B T)^2 W''(\mu) + \mathcal{O}(T^4)$$

To obtain  $\mu$ , we need to invert this expression. However, this tends to be very difficult to invert, especially if we want to include higher order terms. Luckily, we can still obtain an expression for  $\mu$  *even without knowing*  $W$ .

The solution is to write  $\mu$  as a series expansion in the temperature, then *expand again* for small  $T$ . Matching coefficients, we can obtain our result.

$$\mu = \varepsilon_F + \mu_1 k_B T + \mu_2 (k_B T)^2 + \mathcal{O}(T^3)$$

Plugging this ansatz into the equation we wish to invert,

$$\begin{aligned} N &= W(\varepsilon_F + \mu_1 k_B T + \dots) + \frac{\pi^2}{6}(k_B T)^2 W''(\varepsilon_F + \mu_1 k_B T + \dots) + \mathcal{O}(T^2) \\ &= W(\varepsilon_F) + W'(\varepsilon_F)[\mu_1 k_B T + \mu_2 (k_B T)^2 + \dots] + \frac{1}{2}W''(\varepsilon_F)[\mu_1 k_B T + \mu_2 (k_B T)^2 + \dots]^2 + \\ &\quad + \frac{\pi^2}{6}(k_B T)^2 W''(\varepsilon_F) + \frac{\pi^2}{6}(k_B T)^2 W'''(\varepsilon_F)[\mu_1 k_B T + \mu_2 (k_B T)^2 + \dots] + \dots \end{aligned}$$

We recognize  $W(\varepsilon_F) = N$ . Cancelling this, and collecting terms,

$$0 = k_B T [\mu_1 W'(\varepsilon_F)] + (k_B T)^2 \left[ \mu_2 W'(\varepsilon_F) + \frac{1}{2} \mu_1^2 W''(\varepsilon_F) + \frac{\pi^2}{6} W'''(\varepsilon_F) \right] + \mathcal{O}(T^3)$$

We can then solve progressively for the coefficients  $\mu_n$  by setting the coefficient of  $(k_B T)^n$  to zero. Assuming the density of states doesn't vanish near the Fermi energy, we obtain

$$\begin{aligned} \mu_1 &= 0 \\ \mu_2 &= -\frac{\pi^2}{6} \frac{W'''(\varepsilon_F)}{W'(\varepsilon_F)} \\ &\vdots \end{aligned}$$

Or, in terms of the density of states,

$$\mu_2 = -\frac{\pi^2}{6} \frac{D'(\varepsilon_F)}{D(\varepsilon_F)}$$

Thus, we can write

$$\mu(T, N) = \varepsilon_F - \frac{\pi^2}{6} \frac{D'(\varepsilon_F)}{D(\varepsilon_F)} (k_B T)^2 + \mathcal{O}(T^3) \quad (28.9)$$

where  $\varepsilon_F = \varepsilon_F(N)$ ,  $N$  is held constant, and for small temperature. Thus, we see that the chemical potential is approximately quadratic near  $T = 0$ , and is completely determined by the shape of the density of states *at the Fermi Energy*<sup>3</sup>.

We can now insert this into the distribution function, yielding

$$\begin{aligned} f_+(\varepsilon - \mu(T, N)) &\approx \Theta \left( \varepsilon_F - \frac{\pi^2}{6} (k_B T)^2 \frac{D'}{D} - \varepsilon \right) 0 \frac{\pi^2}{6} (k_B T)^2 \delta \left( \varepsilon - \varepsilon_F + \frac{\pi^2}{6} (k_B T)^2 \frac{D'}{D} \right) \\ &\approx \Theta(\varepsilon_F - \varepsilon) + \Theta'(\varepsilon_F - \varepsilon) \left( -\frac{\pi^2}{6} (k_B T)^2 \frac{D'}{D} \right) - \frac{\pi^2}{6} (k_B T)^2 \delta'(\varepsilon - \varepsilon_F) + \mathcal{O}(T^3) \\ &= \Theta(\varepsilon_F - \varepsilon) - \frac{\pi^2}{6} (k_B T)^2 \left[ \frac{D'(\varepsilon_F)}{D(\varepsilon_F)} \delta(\varepsilon - \varepsilon_F) + \delta'(\varepsilon - \varepsilon_F) \right] \end{aligned}$$

which is the low-temperature expansion of the Fermi distribution function at fixed  $N$ . With these results, we are now equipped to examine the low-temperature properties of several thermodynamic observables.

## 28.3 Thermodynamic Observables

We can now calculate the Grand Potential as

$$\Omega(T, \mu) = - \int d\varepsilon W(\varepsilon) f_+(\varepsilon - \mu)$$

---

<sup>3</sup>This general strategy of obtaining an inverse is incredibly useful, especially as these Taylor series can be computed entirely within Mathematica.

$$\approx - \int_{-\infty}^{\mu} d\varepsilon W(\varepsilon) - \frac{\pi^2}{6} (k_B T)^2 D(\mu) + \mathcal{O}(T^4)$$

The free energy is given as the legendre transform

$$\begin{aligned} F(T, N) &= \max_{\mu} \{ \Omega(T, \mu) + \mu N \} = \Omega(T, \mu(T, N)) + N\mu(T, N) \\ &\approx - \int_{-\infty}^{\mu(T, N)} d\varepsilon W(\varepsilon) - \frac{\pi^2}{6} D(\mu(T, N)) + N\mu(T, N) \\ &\approx - \int_{-\infty}^{\mu(T, N)} d\varepsilon W(\varepsilon) - \frac{\pi^2}{6} (k_B T)^2 D(\varepsilon_F) + N \left[ \varepsilon_F - \frac{\pi^2}{6} (k_B T)^2 \frac{D'}{D} \right] + \mathcal{O}(T^3) \\ &= - \left( \int_{-\infty}^{\varepsilon_F} d\varepsilon W(\varepsilon) \right) + \underbrace{W(\varepsilon_F)}_{=N} \frac{\pi^2}{6} \frac{D'}{D} (k_B T)^2 - \frac{\pi^2}{6} (k_B T)^2 D(\varepsilon_F) \\ &\quad + N \left[ \varepsilon_F - \frac{\pi^2}{6} (k_B T)^2 \frac{D'}{D} \right] + \mathcal{O}(T^3) \\ &= - \left( \int_{-\infty}^{\varepsilon_F} d\varepsilon \int_{-\infty}^{\varepsilon} d\varepsilon' D(\varepsilon') \right) - \frac{\pi^2}{6} (k_B T)^2 D(\varepsilon_F) + N\varepsilon_F \end{aligned}$$

flipping the order integration,

$$\begin{aligned} &= - \left( \int_{-\infty}^{\varepsilon_F} d\varepsilon' \int_{\varepsilon'}^{\varepsilon_F} d\varepsilon D(\varepsilon') \right) - \frac{\pi^2}{6} (k_B T)^2 D(\varepsilon_F) + N\varepsilon_F \\ &= - \left( \int_{-\infty}^{\varepsilon_F} d\varepsilon' (\varepsilon_F - \varepsilon') D(\varepsilon') \right) - \frac{\pi^2}{6} (k_B T)^2 D(\varepsilon_F) + N\varepsilon_F \\ &= -(\varepsilon_F N - U_0) - \frac{\pi^2}{6} (k_B T)^2 D(\varepsilon_F) + N\varepsilon_F \\ &= U_0 - \frac{\pi^2}{6} (k_B T)^2 D(\varepsilon_F) + \mathcal{O}(T^4) \end{aligned}$$

and so, the free energy is the ground state energy with a quadratic correction. The entropy can be found easily by

$$S(T, N) = - \left( \frac{\partial F}{\partial T} \right)_{V, N} = \frac{\pi^2}{3} k_B^2 T D(\varepsilon_F) + \mathcal{O}(T^3)$$

The energy is simply a Legendre transform,

$$U(T, N) = F(T, N) + TS(T, N) = U_0 + \frac{\pi^2}{6} (k_B T)^2 D(\varepsilon_F)$$

the heat capacity is a derivative,

$$C_v(T, N) = \frac{\partial U(T, N)}{\partial T} = \frac{\pi^2}{3} k_B^2 T D(\varepsilon_F) + \mathcal{O}(T^3)$$

Of course, the volume dependence is hidden inside  $D(\varepsilon_F)$ . Interestingly, to first order, our heat capacity goes linearly with temperature; contrast this to the bosons we have studied

(namely, the phonon and photon gases), where the low temperature heat capacity goes with  $\sim T^3$ . Thus, for metals, we observe the linear order in the electron heat capacity to dominate the cubic heat capacity for metals. From this slope, we can even measure the density of states at the fermi energy (though, this is a very costly method to do so).

### Why Linear $C_V$ ?

Recall the distribution function,

$$f_+(\varepsilon - \mu) = \frac{1}{e^{\beta(\varepsilon - \mu)} + 1}$$

We have that the interval over which the fermi occupation changes over a range of  $\sim k_B T$ . If we wish to increase the energy of the system, we need to move particles from below the fermi energy to above the fermi energy; this energy change is also on the order of  $k_B T$  (The tail above  $\varepsilon > \mu$  comes from the missing region below  $\varepsilon < \mu$ ). The number of particles is given roughly by the triangular region of the tail, given  $\sim \frac{1}{2} k_B T \cdot \frac{1}{2}$ . Thus, we have  $N \sim k_B T$  particles raised in energy by  $\Delta u \sim k_B T$ , and thus,  $\Delta U \sim (k_B T)^2$  and  $C_V \sim k_B T$ . The density of states in the prefactor comes from weighting the “triangular” tail, and how many of the states are *actually* occupied.





# Chapter 29

## Insulators and Semiconductors

Loosely speaking, this discussion is not so much statistical mechanics, but more so quantum mechanics and solid state physics. We wish to consider the theory of electrons in a solid, but this is its own vast field of research; however, we can give a brief study of crystalline solids, and examine the constraints the crystallinity imposes. This symmetry results in what is known as the *Bloch Theorem*.

### 29.1 Bloch Theorem

Consider electrons in a potential with bravais lattice symmetry. This lattice has so-called lattice basis vectors  $\{\hat{e}_\alpha\}, \alpha \in \{1, 2, 3\}$  with which we can construct arbitrary lattice vectors  $\mathbf{R}_\mathbf{n} = n^\alpha \hat{e}_\alpha, n^\alpha \in \mathbb{Z}$  (using einstein summation notation), which gives us an invariant translation vector. If we then have a hamiltonian

$$H = \frac{p^2}{2m} + U(\mathbf{r})$$

with a periodic potential  $U(\mathbf{r} + \mathbf{R}_\mathbf{n}) = U(\mathbf{r})$ , then the Bloch Theorem states that the eigenfunctions of the hamiltonian can be written as *bloch waves* in the form

$$\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k} \cdot \mathbf{r}} u_{\mathbf{k}}(\mathbf{r}) \tag{29.1}$$

with the following properties:

1.  $u_{\mathbf{k}}$  has lattice symmetry:  $u_{\mathbf{k}}(\mathbf{r} + \mathbf{R}_\mathbf{n}) = u_{\mathbf{k}}(\mathbf{r})$
2. The bloch wave is dependent on a plane wave with lattice momentum  $\mathbf{k}$ .

we can prove this as follows. First, define the reciprocal basis vectors  $\hat{e}^\beta$  such that

$$\hat{e}^\beta \cdot \hat{e}_\alpha = 2\pi\delta_\alpha^\beta$$

We need to define this reciprocal basis as our original lattice basis set is not necessarily orthonormal. Next, we define the translation operator in position space  $T_{\mathbf{R}_n}$  such that the very obvious property is true:

$$T_{\mathbf{R}_n}f(\mathbf{r}) = f(\mathbf{r} + \mathbf{R}_n)$$

Obviously, we have the property

$$T_{\mathbf{R}+\mathbf{R}'} = T_{\mathbf{R}}T_{\mathbf{R}'} = T_{\mathbf{R}'}T_{\mathbf{R}} = T_{\mathbf{R}'+\mathbf{R}}$$

so all translations commute with each other (note that this would *not* be true if we had use rotations  $R_\theta$  rather than translations). Further, this implies that there is a functional relation satisfied by the translation operator, which implies that the eigenvalues of the translation operator satisfy the exponential function's relation

$$\exp[x]\exp[y] = \exp[x+y]$$

or,

$$T_{\mathbf{R}+\mathbf{R}'}|\psi\rangle = c_{\mathbf{R}+\mathbf{R}'}|\psi\rangle = c_{\mathbf{R}}c_{\mathbf{R}'}\psi$$

so

$$c_{\mathbf{R}}c_{\mathbf{R}'} = c_{\mathbf{R}+\mathbf{R}'}$$

or the eigenvalue is an exponential. This follows from simultaneous diagonalization. Further, because

$$T_{\mathbf{R}}U = U \quad \forall \mathbf{R}$$

we have the following:

$$[H, T_{\mathbf{R}}] = [p^2, T_{\mathbf{R}}] = 0$$

because, as we recall, momentum is the generator of translation and necessarily commutes (alternatively, the momentum is independent of position and thus necessarily commutes). Because the hamiltonian commutes with translation operators, we can simultaneously diagonalize them; thus, the eigenvectors of the hamiltonian (eigenstates) are necessarily eigenvectors of the arbitrary translation operators. We thus must have

$$T_{\mathbf{R}_n}\psi(\mathbf{r}) = \psi(\mathbf{r} + \mathbf{R}_n) = c_{\mathbf{n}}\psi(\mathbf{r}) \equiv e^{2\pi i n^\alpha \vartheta_\alpha}\psi(\mathbf{r})$$

We define  $\mathbf{k} = \vartheta_\beta \hat{e}^\beta$ , so then

$$\mathbf{k} \cdot \hat{e}_\alpha = \vartheta_\beta \hat{e}^\beta \cdot \hat{e}_\alpha = 2\pi \vartheta_\beta \delta_\alpha^\beta = 2\pi \vartheta_\alpha$$

so we can rewrite our eigenstate using  $2\pi i n^\alpha \vartheta_\alpha = i n^\alpha k \cdot \hat{e}_\alpha = i \mathbf{k} \cdot \mathbf{R}_n$ . Thus,

$$T_{\mathbf{R}_n}\psi(\mathbf{r}) = e^{i \mathbf{k} \cdot \mathbf{R}_n}\psi(\mathbf{r})$$

Clearly, the vector  $\mathbf{R}_n$  comes from the operator, but  $\mathbf{k}$  comes from  $\psi$ . Thus, we use it to label the eigenfunction,  $\psi = \psi_{\mathbf{k}}$ . Now, we define  $u_{\mathbf{k}}(\mathbf{r}) = e^{-i \mathbf{k} \cdot \mathbf{r}}\psi_{\mathbf{k}}(\mathbf{r})$ . When we apply the translation operator and use our eigenvalue property for  $\psi_{\mathbf{k}}$ ,

$$T_{\mathbf{R}_n}u_{\mathbf{k}}(\mathbf{r}) = e^{-i \mathbf{k} \cdot (\mathbf{r} + \mathbf{R}_n)}\psi_{\mathbf{k}}(\mathbf{r} + \mathbf{R}_n) = e^{i \mathbf{k} \cdot \mathbf{r}}\psi_{\mathbf{k}}(\mathbf{r}) = u_{\mathbf{k}}$$

so  $u_{\mathbf{k}}$  is also an eigenfunction of  $T_{\mathbf{R}_n}$ , but more importantly has lattice symmetry. Inverting our definition for  $u$ , we see that our eigenstates are given

$$\psi_{\mathbf{k}}(\mathbf{r}) = e^{i \mathbf{k} \cdot \mathbf{r}}u_{\mathbf{k}}(\mathbf{r})$$

as desired.

## 29.2 Schrödinger Equation

Now that we have the form of our wavefunction, we can obtain the energy spectrum of our hamiltonian

$$H\psi_{\mathbf{k}} = \left[ \frac{P^2}{2m} + U(r) \right] e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}} \varepsilon_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}}$$

Left multiplying by the conjugate of the plane wave

$$\left[ e^{-i\mathbf{k}\cdot\mathbf{r}} \frac{P^2}{2m} e^{i\mathbf{k}\cdot\mathbf{r}} + U \right] u_{\mathbf{k}} = \varepsilon_{\mathbf{k}} u_{\mathbf{k}}$$

Note that because the plane wave is a function of the position operator, it does not commute with a function of the momentum, and thus we cannot cancel out the exponential. The first term is annoying to compute directly. However, we can consider a new operator

$$\begin{aligned} (-\nabla + \mathbf{k})^2 f &= (-i\nabla + \mathbf{k})(-i\nabla f + \mathbf{k}f) \\ &= -\nabla^2 f - 2i\mathbf{k} \cdot \nabla f + k^2 f \end{aligned}$$

We can compare this operator to the operator we are interested in

$$\begin{aligned} e^{-i\mathbf{k}\cdot\mathbf{r}} (-i\nabla)^2 e^{i\mathbf{k}\cdot\mathbf{r}} f &= e^{-i\mathbf{k}\cdot\mathbf{r}} (-i\nabla)(-i) [i\mathbf{k}e^{i\mathbf{k}\cdot\mathbf{r}} f + e^{i\mathbf{k}\cdot\mathbf{r}} \nabla f] \\ &= e^{-i\mathbf{k}\cdot\mathbf{r}} (-1) [i\mathbf{k} \cdot (i\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} f + i\mathbf{k} \cdot e^{i\mathbf{k}\cdot\mathbf{r}} \nabla f + ik \cdot e^{i\mathbf{k}\cdot\mathbf{r}} \nabla f + e^{i\mathbf{k}\cdot\mathbf{r}} \nabla^2 f] \\ &= k^2 f - 2i\mathbf{k} \cdot \nabla f - \nabla^2 f \end{aligned}$$

which gives us the same result. Because the test function  $f$  was arbitrary, we thus have

$$\frac{(-i\hbar\nabla + \hbar\mathbf{k})^2}{2m} = \frac{\hbar^2(-i\nabla + \mathbf{k})^2}{2m} = e^{-i\mathbf{k}\cdot\mathbf{r}} \frac{P^2}{2m} e^{i\mathbf{k}\cdot\mathbf{r}}$$

so we can rewrite the schrodinger equation

$$\left[ \frac{(\mathbf{P} + \hbar\mathbf{k})^2}{2m} + U(\mathbf{r}) \right] u_{\mathbf{k}}(\mathbf{r}) = \varepsilon_{\mathbf{k}} u_{\mathbf{k}}(\mathbf{r}) \quad (29.2)$$

and so our index  $\mathbf{k}$  now appears in our schrodinger equation. Interestingly, while  $u_{\mathbf{k}}$  has periodicity in real space, we have that  $\psi_{\mathbf{k}}$  has periodicity in *Fourier space*:  $\psi_{\mathbf{k}+\mathbf{G}} = \psi_{\mathbf{k}}$  where  $\mathbf{G}$  denotes a reciprocal lattice vector. We can show this as follows.

### 29.2.1 Reciprocal Space

Let us return to the Bloch theorem and expand  $\psi$  in terms of Fourier modes

$$\psi(\mathbf{r}) = \sum_{\mathbf{q}} c_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}}$$

when we now translate this as an eigenvector of a lattice translation  $T_{\mathbf{R}}$ , we obtain

$$\sum_{\mathbf{q}} c_{\mathbf{q}} e^{i\mathbf{q}\cdot(\mathbf{r}+\mathbf{R})} = \lambda_{\mathbf{R}} \sum_{\mathbf{q}} c_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}}$$

or,

$$0 = \sum_{\mathbf{q}} c_{\mathbf{q}} (\lambda_{\mathbf{R}} - e^{i\mathbf{q} \cdot \mathbf{R}}) e^{i\mathbf{q} \cdot \mathbf{r}} = 0$$

Recognizing  $e^{i\mathbf{q} \cdot \mathbf{r}}$  as a basis of fourier space, we can match coefficients with the RHS, and note that we must have

$$0 = \lambda_{\mathbf{R}} - e^{i\mathbf{q} \cdot \mathbf{R}}$$

for any vector  $\mathbf{q}$  in reciprocal space. Thus, we must have

$$\mathbf{q} \cdot \mathbf{R} = \mathbf{k} \cdot \mathbf{R} + 2\pi n$$

for some reciprocal space vector  $\mathbf{k}$ . Thus, we must be able to write

$$\mathbf{q} = \mathbf{k} + \mathbf{G}$$

for an arbitrary reciprocal lattice vector  $\mathbf{G}$ , where

$$\mathbf{G} \cdot \mathbf{R} = 2\pi n$$

Basically, we have  $\mathbf{q} = \frac{2\pi}{L} \mathbf{n}$  where  $L$  is the system size, while  $\mathbf{G} = \frac{2\pi}{a} \mathbf{n}$  where  $a$  is a lattice constant. In general,  $a < L$  so  $\mathbf{q}$  are more dense than  $\mathbf{G}$ , and  $\{\mathbf{G}\} \in \{\mathbf{q}\}$ , because the system should contain a whole-number of unit cells. Thus,

$$\lambda_{\mathbf{R}} = e^{i\mathbf{q} \cdot \mathbf{R}} = e^{i(\mathbf{k} + \mathbf{G}) \cdot \mathbf{R}} = e^{i\mathbf{k} \cdot \mathbf{R}}$$

Thus,

$$T_{\mathbf{R}} \psi(\mathbf{r}) = e^{i\mathbf{k} \cdot \mathbf{R}} \psi(\mathbf{r})$$

and thus, the basis eigenfunctions are a plane waves with momentum  $\mathbf{k} + \mathbf{G}$ :

$$\psi = e^{i(\mathbf{k} + \mathbf{G}) \cdot \mathbf{r}}$$

$$T_{\mathbf{R}} \psi = \psi(\mathbf{r}) e^{i(\mathbf{k} + \mathbf{G}) \cdot \mathbf{R}} = e^{i\mathbf{k} \cdot \mathbf{R}} \psi(\mathbf{r})$$

### 29.2.2 Energy spectrum

Thus, we can expand any eigenfunction  $\psi_{\mathbf{k}}$  as a Fourier expansion over reciprocal lattice vectors:

$$\psi_{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}'} c_{\mathbf{k} + \mathbf{G}'} e^{i(\mathbf{k} + \mathbf{G}') \cdot \mathbf{r}} \quad (29.3)$$

Pulling out a common factor  $e^{i\mathbf{k} \cdot \mathbf{r}}$ , we see that

$$\psi = e^{i\mathbf{k} \cdot \mathbf{r}} \sum_{\mathbf{G}'} c_{\mathbf{k} + \mathbf{G}'} e^{i\mathbf{G}' \cdot \mathbf{r}}$$

we interpret this second term as

$$u_{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}'} c_{\mathbf{k} + \mathbf{G}'} e^{i\mathbf{G}' \cdot \mathbf{r}} \quad (29.4)$$

Further, we can show that this representation of  $u_{\mathbf{k}}(\mathbf{r})$  has lattice symmetry:

$$\begin{aligned} T_{\mathbf{R}}u_{\mathbf{k}}(\mathbf{r}) &= u_{\mathbf{k}}(\mathbf{r} + \mathbf{R}) \\ &= \sum_{\mathbf{G}'} c_{\mathbf{k}+\mathbf{G}'} e^{i\mathbf{G}' \cdot (\mathbf{r} + \mathbf{R})} \\ &= \sum_{\mathbf{G}'} c_{\mathbf{k}+\mathbf{G}'} e^{i\mathbf{G}' \cdot \mathbf{r}} e^{i\mathbf{G}' \cdot \mathbf{R}} \end{aligned}$$

Recognizing  $e^{i\mathbf{G}' \cdot \mathbf{R}} = e^{i2\pi n}$ , we see trivially, that

$$\begin{aligned} &= \sum_{\mathbf{G}'} c_{\mathbf{k}+\mathbf{G}'} e^{i\mathbf{G}' \cdot \mathbf{r}} \\ &= u_{\mathbf{k}}(\mathbf{r}) \end{aligned}$$

Another consequence of our representation is a description of how our wavefunctions transform when shifted in  $\mathbf{k}$  space. Consider a new function,  $u_{\mathbf{k}+\mathbf{G}}$  such that  $\mathbf{G}' = \mathbf{G}'' - \mathbf{G}$ . Further, summing over all  $\mathbf{G}'$  is the same as summing over all  $\mathbf{G}''$ . Thus, we can write

$$\begin{aligned} u_{\mathbf{k}+\mathbf{G}}(\mathbf{r}) &= \sum_{\mathbf{G}'} u_{(\mathbf{k}+\mathbf{G})+\mathbf{G}'} e^{i\mathbf{G}' \cdot \mathbf{r}} \\ &= \sum_{\mathbf{G}''} c_{\mathbf{k}+\mathbf{G}''} e^{i(\mathbf{G}'' - \mathbf{G}) \cdot \mathbf{r}} \\ &= e^{-i\mathbf{G} \cdot \mathbf{r}} \sum_{\mathbf{G}''} c_{\mathbf{k}+\mathbf{G}''} e^{i\mathbf{G}'' \cdot \mathbf{r}} \\ &= e^{-i\mathbf{G} \cdot \mathbf{r}} u_{\mathbf{k}}(\mathbf{r}) \end{aligned}$$

Thus, in our wavefunction, we see that

$$\begin{aligned} \psi_{\mathbf{k}+\mathbf{G}}(\mathbf{r}) &= e^{i(\mathbf{k}+\mathbf{G}) \cdot \mathbf{r}} u_{\mathbf{k}+\mathbf{G}}(\mathbf{r}) \\ &= e^{i\mathbf{k} \cdot \mathbf{r}} u_{\mathbf{k}}(\mathbf{r}) \\ &= \psi_{\mathbf{k}}(\mathbf{r}) \end{aligned}$$

Thus, we have shown  $\psi_{\mathbf{k}}(\mathbf{r})$  has reciprocal lattice symmetry.

### 29.2.3 Brillouin Zone Folding

Since  $\psi_{\mathbf{k}}(\mathbf{r})$  has reciprocal lattice periodicity, then so do the eigenvalues. Thus,

$$\varepsilon_{\mathbf{k}+\mathbf{G}} = \varepsilon_{\mathbf{k}} \tag{29.5}$$

This result gives rise to “folding” of the spectrum into the First Brillouin zone.

While solving the Schrödinger equation for a given lattice potential is generally very hard, we can consider a so-called *empty lattice*,

$$U(\mathbf{r}) = 0$$

This allows us to consider what the lattice symmetry *alone* enforces, and not the shape of the potential itself.

Trivially, our Schrödinger equation becomes

$$\frac{(\mathbf{P} + \hbar\mathbf{k})^2}{2m}u_{\mathbf{k}}(\mathbf{r}) = \varepsilon_{\mathbf{k}}u_{\mathbf{k}}(\mathbf{r})$$

Trivially, this is solved by  $u_{\mathbf{k}} = a_{\mathbf{k}}$  for some constant  $a$ . This has the eigenvalues  $\varepsilon_{\mathbf{k}} = \frac{\hbar^2 k^2}{2m}$ , as expected of a free particle.

{diagram}

However, we must enforce symmetry on the eigenvalue. We do this by *folding* the dispersion relation at the Brillouin boundary. We enforce lattice symmetry by replicating the dispersion relation at every reciprocal lattice vector  $\mathbf{G}$ .

{diagram}

Then, we consider only the first Brillouin zone,

{diagram}

We see that this dispersion relation is the same as if we were to fold the original, solitary parabolic dispersion relation every time it intersects a brillouin boundary.

If we were to include interactions via perturbation theory, we see that any intersections (i.e. where two parabolas intersect, or where the dispersion relation meets the brillouin boundary), the interaction causes the degenerate eigenvalues to split into a higher and a lower band

{diagram}

Thus, on the vertical axis, we have discrete regions where we have valid energies—these are known as *bands*. In between the bands, where there are no valid energies, we have *gaps*. In principle, we can use the transformation theorem to determine the density of states for this dispersion relation. However, in practice, this is very difficult, and so we will only consider the bands qualitatively.

As we add electrons, lower bands get filled and higher bands remain empty, Recall the Fermi energy, which is exactly in the middle of the gap between filled and empty states. If we increase the temperature, then electrons in the filled, *valence band* will get promoted into the empty, *conduction band*. The dispersion relation at this band gap looks surprisingly like the relativistic dispersion relation

$$E = \pm\sqrt{p^2c^2 + m^2c^4}$$

In the vicinity of these edge states, we expect the dispersion relation to be approximated by parabolas—near the band gap, the dispersion relations look like those for free particles.

## 29.3 Semiconductor Band Structure

We will use the model density of states given

$$D(\varepsilon) = \begin{cases} A(\varepsilon - \varepsilon_c)^a & \varepsilon_c < \varepsilon \\ 0 & \varepsilon_v < \varepsilon < \varepsilon_c \\ B(\varepsilon_v - \varepsilon)^b & \varepsilon < \varepsilon_v \end{cases} \quad (29.6)$$

for some constants  $a, A, b, B$ . Of course,  $\varepsilon_c$  is the lowest state in the conduction band and  $\varepsilon_v$  is the highest state in the valence band. Let us assume that the valence band is completely full and the conduction band is completely empty at  $T = 0$ . By the rule of thumb, we trivially have

$$\varepsilon_F = \frac{\varepsilon_c + \varepsilon_v}{2}$$

however, we can verify this. We have

$$\begin{aligned} N &= \left[ \int_{-\infty}^{\varepsilon_v} d\varepsilon + \int_{\varepsilon_c}^{\infty} \right] D(\varepsilon) f_+(\varepsilon - \mu) \\ &= \underbrace{\int_{-\infty}^{\varepsilon_v} d\varepsilon D(\varepsilon)}_N - \underbrace{\int_{-\infty}^{\varepsilon_v} d\varepsilon D(\varepsilon) [1 - f_+(\varepsilon - \mu)]}_{N_h} + \underbrace{\int_{\varepsilon_c}^{\infty} d\varepsilon D(\varepsilon) f_+(\varepsilon - \mu)}_{N_c} \end{aligned}$$

Where we recognize the factor  $[1 - f_+]$  as the probability that a state is *unoccupied*, or is a hole. Thus, the second term becomes the number of holes, and the third term is the number of charge carriers in the conduction band. Note that we can substitute the first term for  $N$  because  $D(\varepsilon)$  for  $\varepsilon \in [\varepsilon_v, \varepsilon_F] = 0$ . Thus,

$$N_h = N_c$$

We can compute these by

$$\begin{aligned} N_h &= \int_{-\infty}^{\varepsilon_v} d\varepsilon' D(\varepsilon') \frac{1}{e^{-\beta(\varepsilon' - \mu)} + 1} = \int_0^{\infty} d\varepsilon D(\varepsilon_v - \varepsilon) \frac{1}{e^{\beta(\varepsilon + \mu - \varepsilon_v)} + 1} \\ N_c &= \int_{\varepsilon_c}^{\infty} d\varepsilon' D(\varepsilon') \frac{1}{e^{\beta(\varepsilon' - \mu)} + 1} = \int_0^{\infty} d\varepsilon D(\varepsilon + \varepsilon_c) \frac{1}{e^{\beta(\varepsilon + \varepsilon_c - \mu)} + 1} \end{aligned}$$

where in the first equation we substitute  $\varepsilon = \varepsilon_v - \varepsilon'$  and in the second  $\varepsilon = \varepsilon' - \varepsilon_c$ . We can then approximate the value of these integrals as follows. First, assume that  $\mu \approx \varepsilon_F$  and that this is roughly independent of  $T$ . Second, we assume that  $k_B T \ll \frac{\varepsilon_v - \varepsilon_c}{2} \approx \mu - \varepsilon_v \approx \varepsilon_c - \mu$ . In this case, we can neglect the 1 in both integrals. Thus,

$$\begin{aligned} N_h &\approx e^{-\beta(\mu - \varepsilon_v)} \int_0^{\infty} d\varepsilon D(\varepsilon_v - \varepsilon) e^{-\beta\varepsilon} \equiv a(T) e^{-\beta(\mu - \varepsilon_v)} \\ N_c &\approx e^{-\beta(\varepsilon_c - \mu)} \int_0^{\infty} d\varepsilon D(\varepsilon + \varepsilon_c) e^{-\beta\varepsilon} \equiv b(T) e^{-\beta(\varepsilon_c - \mu)} \end{aligned}$$

where we see that  $a, b$  are *Laplace Transforms* of the density of states.

We know that we have  $N_h = N_c$ , so we can find the chemical potential.

$$ae^{-\beta(\mu-\varepsilon_v)} = be^{-\beta(\varepsilon_c-\mu)}$$

or

$$\mu(T) = \frac{\varepsilon_c + \varepsilon_v}{2} + \frac{1}{2}k_B T \ln \frac{a(T)}{b(T)}$$

If the band is symmetric about the fermi energy, then we see that  $a = b$  and thus the chemical potential is equally the fermi energy.

Finally, for conductivity, we are interested in the number of charge carriers,  $N_c = N_h = \sqrt{N_c N_h}$ . This final expression is useful in that *it does not require us to know the chemical potential*:

$$\sqrt{N_h N_c} = e^{-\beta(\varepsilon_c - \varepsilon_v)/2} \sqrt{ab} = e^{-\beta\varepsilon_g/2} \sqrt{ab}$$

For semiconductors, we have the band gap  $\varepsilon_g \sim 1 \text{ eV}$ , and at RT  $\beta \sim 1/40 \text{ eV}$ , so  $e^{-\beta\varepsilon_g/2}$  is very small, and thus the number of charge carriers is very small. One may argue that the term  $\sqrt{ab}$  may increase  $N_c$  sufficiently to allow conduction; however, this is not the case. Often, the band energies are such that the dispersion relation forms two opposing parabolas at the band gap. In the conduction band, we have

$$\varepsilon(\mathbf{k}) = \varepsilon_c + \frac{\hbar^2 k^2}{2m_c}$$

and in the valence band

$$\varepsilon(\mathbf{k}) = \varepsilon_v - \frac{\hbar^2 k^2}{2m_v}$$

for effective masses  $m_c, m_v$ . The shape of the band—in particular, the curvature of the band due to perturbative splitting of the degenerate energies—determines this effective mass. Thus, our particles look like free particles, and we know the density of states for a free particle. Fix  $d = 3$ . Then,

$$D(\varepsilon) = \frac{V}{2\pi^2} \left[ \left( \frac{2m_c}{\hbar^2} \right)^{3/2} \Theta(\varepsilon - \varepsilon_c) \sqrt{\varepsilon - \varepsilon_c} + \left( \frac{2m_v}{\hbar^2} \right)^{3/2} \Theta(\varepsilon_v - \varepsilon) \sqrt{\varepsilon_v - \varepsilon} \right]$$

Where the step function  $\Theta(x)$  is zero when  $x < 0$  and one when  $x \geq 0$ , and the densities of state are shifted according to the band offsets,  $\varepsilon_v, \varepsilon_c$ . Further, recall that we care only for  $D(\varepsilon)$  for  $\varepsilon \sim \varepsilon_F$ , as we are dealing with a fermi distribution, so it doesn't matter if our assumptions stop holding far away from the fermi energy. Then,

$$b(T) = \frac{V}{2\pi^2} \left( \frac{2m_c}{\hbar^2} \right)^{3/2} \int_0^\infty d\varepsilon \sqrt{\varepsilon} e^{-\beta\varepsilon} = \frac{1}{4} V \left( \frac{2m_c}{\pi\hbar^2} \right)^{3/2} (k_B T)^{3/2}$$

Similarly,

$$a(T) = \frac{1}{4} V \left( \frac{2m_v}{\pi\hbar^2} \right)^{3/2} (k_B T)^{3/2}$$



and so,

$$N_h = \frac{1}{4}V \left( \frac{2m_v}{\pi\hbar^2} \right)^{3/2} (k_B T)^{3/2} e^{-\beta(\mu - \varepsilon_v)}$$

Further,

$$N_c = N_h = \sqrt{N_c N_h} = e^{-\beta(\varepsilon_c - \varepsilon_v)} \frac{1}{4}V \left( \frac{2k_B T}{\pi\hbar^2} \right)^{3/2} (m_c m_v)^{3/4}$$

It is important to note the factor of  $V$  which indicates that the number of charge carriers is extensive, as desired. The temperature dependence is complicated and

$$N_c \sim T^{3/2} e^{-\varepsilon_g/k_B T}$$

and so, we have strong exponential suppression that often drowns out the  $T^{3/2}$ . This is because the band gap  $\varepsilon_g \sim 1$  eV, the exponential is of order  $e^{-40}$  at standard temperatures. In order to get large numbers of charge carriers, we would need to heat the material to  $\sim 40\times$  room temperature, which is usually unreasonable. Rather, to increase the conductivity, we often rely on *doping*.

## 29.4 Doping

Doping adds a small fraction of atoms which have one more (less) electron than the bulk crystal. For example, we can dope silicon, which has 4 valence electrons, with phosphorus atoms, which has 5 electrons and *n-dopes* the silicon with extra electrons, or with boron, which has 3 valence electrons and *p-dopes* the silicon with extra holes.

When we n-dope a semiconductor, we can view the extra electron as orbiting the extra charge on the phosphorous, in a bohr-like orbit. If we give that electron a high enough energy, it is no longer a bound state and becomes a free particle and can conduct. However, it is not simply a hydrogenic atom—the crystal structure forces the electron mass to be an effective mass. Further, the dielectric constant changes, which alters the coulombic forces within the material. These two changes together within rydberg

$$\text{Ryd} \sim \frac{m}{\varepsilon^2}$$

shrinks the ionization energy from the 13.6 eV to a quantity that is typically much smaller than the band gap. Thus, the electrons given by doping occupy a small valence band much closer to the bulk conduction band.

Interestingly, the donor state is doubly degenerate because it can allot for either spin up or spin down. Thus, it should, in principle, be possible to doubly occupy it. However, due to coulomb repulsion, it can only be singly occupied. The energy is so elevated that this doubly occupied state is unstable at STP. The reason we have not had to discuss coulombic repulsion before is that the other electrons are localized in reciprocal space and delocalized in position space. Because all of the electrons are smeared out, they do not interact with each other outside of as a shielding effect. In contrast, the two electrons in the donor state

are localized in position space and thus interact strongly. Further, this state is not very well-behaved in terms of statistical mechanics.

Consider the grand canonical partition function of that state. If we neglect the coulomb repulsion, we have

$$\sum_{n_{\varepsilon_d}} e^{-\beta(\varepsilon_d - \mu)n_{\varepsilon_d}} = 1 + 2e^{-\beta(\varepsilon_d - \mu)} + e^{-\beta(\varepsilon_d - \mu)^2} = (1 + e^{-\beta(\varepsilon_d - \mu)})^2$$

so the partition function factorizes. However, the last term in the sum is forbidden due to the coulomb repulsion, as it sends the energy way up. Thus,

$$\langle n_{\varepsilon_d} \rangle = \frac{0 \cdot 1 + 1 \cdot 2e^{-\beta(\varepsilon_d - \mu)}}{1 + 2e^{-\beta(\varepsilon_d - \mu)}} = \frac{1}{\frac{1}{2}e^{\beta(\varepsilon_d - \mu)} + 1}$$

which has an extra factor of 1/2 compared to the Fermi distribution. This is sometimes known as “semiconductor statistics.”

### 29.4.1 Semiconductor Statistics Example

Consider a “three level system” with an  $N_v$ -fold degenerate valence level  $\varepsilon_v$ , an  $N_c$ -fold degenerate conduction level  $\varepsilon_c$ , and an  $N_d$ -fold degenerate donor level  $\varepsilon_d$ . We use these condensed “band” levels rather than true bands in order to simplify calculation. Further, we will assume that the valence band is full and conduction band is empty, as customary, and that the temperature is low enough where only transitions from the donor band are significant ( $k_B T \ll \varepsilon_d - \varepsilon_v$ ). Thus, Once again, we will assume that  $k_B T \ll \begin{cases} \mu - \varepsilon_d \\ \varepsilon_c - \mu \end{cases}$ .

Indeed, when we consider the chemical potential, when we use semiconductor statistics, we see that

$$\mu = \frac{\varepsilon_c + \varepsilon_d}{2} + \frac{1}{2}k_B T \log \frac{N_d}{2N_c}$$

however, the low temperature limit shows us the fermi energy is

$$\varepsilon_F = \frac{\varepsilon_c + \varepsilon_d}{2}$$

however, this seems to violate our rule of thumb, as the lowest unoccupied state is *also in the donor band*! This is reconciled when we consider the fact that the unoccupied state in the donor band is unoccupy-able, and so the lowest occupiable state is in the conduction band, and hence the Fermi level is between these two bands.

Thus, we are able to use the same approximation we used for the undoped semiconductor.

$$N = N_d = \frac{N_d}{\frac{1}{2}e^{\beta(\varepsilon_d - \mu)} + 1} + \frac{N_c}{e^{\beta(\varepsilon_c - \mu)} + 1}$$

$$= N_d \left[ 1 - \frac{1}{2e^{-\beta(\varepsilon_d - \mu)} + 1} \right] + \frac{N_c}{e^{\beta(\varepsilon_c - \mu)} + 1}$$

We will find that in n-doped semiconductors, the charge carriers are primarily from the donor band and are negatively charged electrons moving in the conduction band, while in p-doped semiconductors, the charge carriers are primarily positively charged holes moving in the valence band. When we make a PN junction, we want to have the chemical potential (and thus the fermi energy) to be the same (for equilibrium principle) we see that the two fermi energies move toward each other by moving electrons from the N-doped side to the P-doped side, an internal electric field is formed due to charge imbalance. When we add an external potential, the bands further deform until we see the behaviour typical of a diode.



# Chapter 30

## Interactions and the Ising Model

Now that we have completed our discussion of classical, bose, and fermi ideal gases, we will now consider how particle interactions change the thermodynamics of a system. Particle interactions make it so that the partition function doesn't factorize, and thus it is very hard to consider. However, it is these interactions which allow for the interesting phenomena in nature, namely phase transitions.

In particular, consider the thermal equation of state for the van der Waals gas,

$$\left(P + a\frac{N^2}{V^2}\right)(V - bN) = NK_B T$$

We see that in the first term,  $an^2$ , comes from attractions, and lowers the pressure relative to the ideal gas (less pressure is required for the same volume and temperature). In the second term,  $-bN$  comes from short-range repulsions between particles, and thus the accessible volume is smaller than expected, or we need more volume for the same pressure and temperature. A pair potential  $U(r)$  that could lead to this interaction could be something like the morse potential. However, we did not derive this equation based on first principles—we didn't derive this equation from the interaction potential. There is no theory that we can compare predictions with.

In general, there are two ways we can solve systems with interactions<sup>1</sup>. First, of course, we can find an analytical solution. However, there are very few circumstances where this is possible, and even if it is, it is often very difficult to do so. The other way, which is what we will discuss, is to find a clever approximation that allows the system to be “solved.”

These approximations further fall into two classes. First, we have Mean Field Theories, where we ignore the correlated fluctuations—we set every particle into an “effective” potential based on the average interactions of the other particles. This is typically a good first step. However, it cannot be used when such fluctuations matter, like in the vicinity of a critical point (e.g. that of a vdW gas). This is due to the fluctuation-response theorem, as if the response (such as heat capacity, compressibility, etc) diverges, the fluctuations do as well. Under such

---

<sup>1</sup>Of course, we can always brute force numerical partition functions, namely through Monte Carlo and Molecular Dynamics simulations.

circumstances, MFT isn't even a qualitative description. The second class is anything else. Solutions that don't rely on MFT are often harder to calculate and fine-tuned based on the hamiltonian of interest.

## 30.1 Ising Model

To study interacting systems, a wide number of toy models have been developed. While highly idealized, these toy models attempt to capture the “essence” of the problem of interest. In particular, one of the most famous systems is the *Ising model*<sup>2</sup>.

Consider a lattice in  $d$  spatial dimensions. Note in  $d > 1$  we need to specify the connections for the different types of lattice, but a simple cubic lattice is often assumed. On the lattice sites, we have classical degrees of freedom which are denoted “spins,” which can take values<sup>3</sup>  $\sigma \in \{-1, 1\}$ . The configuration of the system is given by fixing a value  $\sigma_i$  on each location  $i$  on the lattice.

There are two ways in which a given configuration (or, *microstate*) of spins contributes to the energy of the system. First, we want to have a function  $j(\sigma_i, \sigma_j)$  such that anti-aligned and align spins have different energies, and another function  $h(\sigma_i)$  which imposes an energy difference between the spin  $-1$  and spin  $1$  states. Naturally, we can create functions over large or different sized clusters, but we will restrict our discussion to nearest neighbour interactions.

In the Ising model, the simplest model functions  $h$  and  $j$  are those which are linear in each of their arguments. In particular, we have

$$h(\sigma_i) = -h \cdot \sigma_i$$

or the spin up energy is  $-h$  and the spin down is  $+h$ , or spin up is favoured. Next, we want to couple the neighbouring spins. The simplest way to do so is to just multiply the two together to see if they are aligned or anti aligned (similar to taking an inner product for a projection). We define this to be

$$j(\sigma_i, \sigma_j) = -J(\sigma_i \cdot \sigma_j)$$

thus, if the spins are aligned, we have  $\sigma_i \cdot \sigma_j = 1$  and the energy is depressed, and vice-versa for anti-aligned spin. Our total hamiltonian is thus given

$$H = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i \quad (30.1)$$

where the brackets  $\langle ij \rangle$  denotes the set of all nearest-neighbour pairs (or bonds). This accounts for the fact that if we simply sum over  $i$  and  $j$  we will double count the bond  $ij$  from the bond  $ji$ . Note that the type of lattice is encoded implicitly within  $\langle ij \rangle$ .

<sup>2</sup>This model was invented by Wilhelm Lenz (1920), who gave it to his PhD student, Ernst Ising. Ising managed to exactly solve the problem in 1D for his thesis (1924).

<sup>3</sup>In CVM, this is extended to accept a much wider range of “spins,” namely different types and configurations of atoms. In this way, we can extend the Ising model to study metal alloys.

Interestingly, the answer to the Ising model depends on what dimension the crystal lattice is, even qualitatively. Similar to how BECs only exist in  $d > 2$ , we will see that there is a dimension-dependent existence of the phase transition. Further, we can actually solve this model exactly in 1D, but the solution is boring. There is also an exact 2D solution when  $h = 0$ , and has very interesting consequences, but this is very complicated. However, there is no known exact solution for 3D, even for the special case  $h = 0$ . Curiously, while we still don't have exact solutions for  $d \geq 4$ , the “essence” of the problem becomes simple again. In particular  $d = 4$  might be interesting when the dimensions aren't strictly spatial, and the solution for  $d = 0$  won a nobel prize.

## 30.2 No Coupling

While it may seem useless to consider the case with no coupling,

$$H = -h \sum_i \sigma_i$$

we will see that it is actually a very important key to approximating the case where the interactions are no longer negligible. We can write our partition function as a sum over the space of configurations  $\{\sigma\}$

$$Z = \sum_{\{\sigma\}} e^{-\beta(-h \sum_i \sigma_i)}$$

Of course, this can be written out as

$$= \sum_{\sigma_1=-1}^1 \cdots \sum_{\sigma_N=-1}^1 e^{\beta h \sigma_1} \cdots e^{\beta h \sigma_N}$$

so we can factorize as

$$\begin{aligned} &= \left( \sum_{\sigma=-1}^1 \right)^N \\ &= (2 \cosh \beta h)^N \end{aligned}$$

Thus, our free energy is given

$$F = -k_B T \log Z = -N k_B T \log(2 \cosh \beta h)$$

Another interesting quantity is the mean magnetization,

$$M = \left\langle \sum_{i=1}^N \sigma_i \right\rangle$$

Averaging over the canonical state, we can find this as

$$\begin{aligned}
 M &= \frac{\sum_{\{\sigma\}} (\sum_i \sigma_i) e^{\beta h \sum_i \sigma_i}}{\sum_{\{\sigma\}} e^{\beta h \sum_i \sigma_i}} \\
 &= \frac{\sum_{\{\sigma\}} \frac{1}{\beta} \frac{\partial}{\partial h} e^{\beta h \sum_i \sigma_i}}{\sum_{\{\sigma\}} e^{\beta h \sum_i \sigma_i}} \\
 &= \frac{\frac{1}{\beta} \frac{\partial}{\partial h} \sum_{\{\sigma\}} e^{\beta h \sum_i \sigma_i}}{\sum_{\{\sigma\}} e^{\beta h \sum_i \sigma_i}} \\
 &= \frac{1}{\beta} \frac{\partial}{\partial h} \ln Z
 \end{aligned}$$

Pulling the  $\beta$  into the derivative, we find that we can write the magnetization as

$$M = -\frac{\partial F}{\partial h} \quad (30.2)$$

thus, we see that the magnetization and field strength are canonically conjugate pairs. Note that even if we include the coupling term, *nothing changes*—this expression for magnetization holds even with coupling. This is why we chose to differentiate wrt  $h$  rather than wrt  $\beta$ .

Applying this formula to the case where there is no coupling, we see that

$$M = -\frac{\partial F}{\partial h} = -\frac{\partial}{\partial h} [-Nk_B T \ln(2 \cosh \beta h)] = Nk_B T \frac{2\beta \sinh \beta h}{2 \cosh \beta h} = N \tanh \beta h$$

or, the magnetization per spin

$$m = \frac{M}{N} = \tanh \beta h$$

The function  $\tanh$  has three limits. For  $x \approx 0$ , it approaches  $\tanh(x) \rightarrow x$ , while at  $x \rightarrow \pm\infty$  it approaches  $\tanh(x) \rightarrow \pm 1$ . At small field strength,  $h/k_B T \ll 1$ , we see that there is then a linear response regime, or  $m \propto h$ . However, there is a limit to the magnetization—each spin can only accomodate magnetization  $m = 1$ . Thus, the response deviates from linearity due to *saturation*.

Importantly, we have when there is zero coupling,  $h = 0 \implies m = 0$ , so the material is not spontaneously magnetic. However, this makes sense, as there is no force that would favour magnetization of any kind; entropic effects dominate the configuration and causes a even, random distribution of spins.

### 30.3 Ising Chain

Let us now consider a spin coupling term

$$H_j = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j$$



which should favour aligned spins. We would then expect there to be spontaneous magnetization to arise<sup>4</sup>. Indeed, this is the energetically favoured state. However, the stability of the system should be given instead by the free energy, which has a contribution from the entropy. When we observe magnetization, the entropy of the system decreases, which is *not* favoured. For example, if we have all spins aligned, there is one microstate, and the entropy is thus given  $S = 0$ . On the other hand, if we add one misaligned spin, then there are  $N$  microstates and  $S = k_B \ln N > 0$ . Thus, we wish to see under which regimes the energy dominates, and which regimes the entropy dominates; we wish to minimise the free energy

$$F = E - TS$$

To figure out this balance, we need to compute the partition function while including interactions. However, as discussed many times before, the interactions make  $Z$  unfactorizable. We will in particular consider the Ising model in  $d = 1$ , while enforcing periodic boundary conditions<sup>5</sup>. Since all of the bonds have to lie in a line, we can denumerate each bond by only its first atom. Thus, our hamiltonian becomes

$$H = -J \sum_{i=1}^N s_i s_{i+1} - h \sum_i s_i$$

To make computation easier, we can rewrite the  $h$  sum to yield

$$H = -J \sum_{i=1}^N s_i s_{i+1} - \frac{1}{2} h \sum_{i=1}^N (s_i + s_{i+1}) \quad (30.3)$$

Our partition function can then be calculated as follows:

$$\begin{aligned} Z &= \sum_{\{s\}} \exp \left[ -\beta \left( -J \sum_{i=1}^N s_i s_{i+1} - \frac{1}{2} h \sum_{i=1}^N (s_i + s_{i+1}) \right) \right] \\ &= \sum_{\{s\}} \prod_{i=1}^N \exp \left[ \beta J s_i s_{i+1} + \frac{1}{2} \beta h (s_i + s_{i+1}) \right] \end{aligned}$$

In his original thesis, Ising used many pages of combinatorics to find every possible configuration to essentially brute force the solution. Obviously, we will not do this. Rather, we will use an incredibly ingenious and elegant solution to this extremely difficult problem<sup>6</sup>. We see that the exponential can be indexed by the values of  $s_i$  and  $s_{i+1}$ , which each can take two values. Thus, we will make the exponential a  $2 \times 2$  matrix  $T_{s_i, s_{i+1}}$  called the *transfer matrix*. WLOG, we will fix  $T_{1,1}$  to be the topright so we can write

$$T = \begin{bmatrix} e^{\beta(J+h)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-h)} \end{bmatrix}$$

Plugging this into our expression for the partition function,

$$= \sum_{\{s\}} \prod_{i=1}^N T_{s_i, s_{i+1}}$$

<sup>4</sup>This is why Lenz gave this problem to Ising to study

<sup>5</sup>Because of the periodic boundary conditions, this is sometimes called the *Ising Ring*.

$$= \sum_{s_1=-1}^1 \sum_{s_2=-1}^1 \cdots \sum_{s_N=-1}^1 T_{s_1, s_2} T_{s_2, s_3} \cdots T_{s_N, s_1}$$

We now see that we are summing over each repeated index, which shows us that we can write this as matrix multiplication!

$$\begin{aligned} &= \sum_{s_1=-1}^1 (T^N)_{s_1, s_1} \\ &= \text{tr}(T^N) \end{aligned}$$

This is particularly useful as the trace is basis-invariant. Trivially,  $T$  is a symmetric, real matrix and so it can be diagonalized. Assume that  $U$  diagonalizes  $T$  such that

$$\tilde{T} = U T U^{-1} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

Using the fact that we can insert the identity anywhere we want, and cyclicly permute,

$$\begin{aligned} \text{tr}(T^N) &= \text{tr}(T I T \cdots T I) = \text{tr}(T U^{-1} U T \cdots T U^{-1} U) \\ &= \text{tr}(U T U^{-1} U T \cdots T U^{-1}) \\ &= \text{tr}(\tilde{T}^N) = \lambda_1^N + \lambda_2^N \end{aligned}$$

*This last equality is our partition function!*

$$Z = \lambda_1^N + \lambda_2^N \tag{30.4}$$

Thus, for the 1D chain, all we need to do is find the eigenvalues of a  $2 \times 2$  matrix! We can do this, of course, from the characteristic polynomial

$$\begin{aligned} 0 &= \det(T - \lambda I) \\ &= (e^{\beta(J+h)} - \lambda)(e^{\beta(J-h)} - \lambda) - e^{-2\beta J} \\ &= e^{2\beta J} - \lambda [e^{\beta(J+h)} + e^{\beta(J-h)}] + \lambda^2 e^{-2\beta J} \\ &= \lambda^2 - \lambda e^{\beta J} 2 \cosh(\beta h) + 2 \sinh(2\beta J) \end{aligned}$$

so

$$\lambda_{\pm} = e^{\beta J} \cosh(\beta h) \pm \sqrt{[e^{\beta J} \cosh(\beta h)]^2 - 2 \sinh(2\beta J)} \tag{30.5}$$

We can now compute the free energy:

$$\begin{aligned} F &= -k_B T \log(\lambda_+^N + \lambda_-^N) \\ &= -k_B T \log \lambda_+^N \left[ 1 + \left( \frac{\lambda_-}{\lambda_+} \right)^N \right] \\ &= -N k_B T \log \lambda_+ - k_B T \log \left[ 1 + \left( \frac{\lambda_-}{\lambda_+} \right)^N \right] \end{aligned}$$

Recognizing that  $\lambda_- < \lambda_+$ , and  $N \rightarrow \infty$  in the thermodynamic limit, we can simplify this by using  $(\lambda_-/\lambda_+)^N \ll 1$  so

$$F \rightarrow -Nk_B T \log \lambda_+ \quad (30.6)$$

This is the *exact, analytical solution* to the 1D Ising chain in the thermodynamic limit. From this expression, we can now find the magnetization. After some tedious algebra, we find

$$m = \frac{1}{N} \left( -\frac{\partial F}{\partial h} \right) = \frac{e^{\beta J} \sinh(\beta h)}{\sqrt{e^{2\beta J} \cosh^2(\beta h) - 2 \sinh(2\beta J)}} \quad (30.7)$$

Under small fields,  $\beta h \rightarrow 0$ , we see that

$$m(h) \rightarrow e^{2\beta J} \beta h - \frac{1}{6} (3e^{6\beta J} - e^{2\beta J}) (\beta h)^3 + O[(\beta h)^5]$$

Being an odd function, this matches our physical expectation (if it were an even function of  $\beta h$ , we'd expect the direction of magnetization to be independent of the direction of the field). Of course, as  $J \rightarrow 0$  we recover our previous expression, where  $m = \tanh(\beta h)$ . As we increase  $\beta J$ , we see that the material becomes easier to magnetize; the linear response  $e^{2\beta J} > 1$  and the spins prefer to be aligned. Consequently, the material saturates at a lower field strength as well. On the other hand, if we decrease  $J$ , we see that the linear response  $e^{-2\beta|J|} < 1$  and so the material resists magnetization. In particular,  $J < 0$  favours spin anti-alignment, and so a greater field needs to be applied to force a bulk magnetization. The case  $J < 0$  is also interesting in that the curvature is no longer monotonic—there are inflection points beside that at the origin.

**{diagram}**

Quite disappointingly, however, the zero-field magnetization is always zero—there is no spontaneous magnetization. This confused Ising for the following reason: a  $d = 3$  crystal can be viewed as a collection of  $d = 1$  chains, which do not spontaneously magnetize, so  $d = 3$  probably shouldn't either. *However, we know that this is not the case*, and so the model was considered a poor model of ferromagnetism.

Luckily, this was a premature conclusion, as the  $d = 1$  case is actually a very special case. To see why, let us revisit our previous entropy argument and consider an open Ising chain. In the absence of a magnetic field, we see that we have two ground states—all spin up or all spin down. Rather than  $2N$  possible excited states, there are only  $2(N - 1)$  states—we choose one of the  $N - 1$  bonds, and invert the rest of the chain. Now, we see that unlike the  $J = 0$  case mentioned earlier, in the  $h = 0$  case, *the magnetization ranges anywhere from  $N - 2$  to  $2 - N$* . Further, compared to the ground state, this leads to the following change in free energy:

$$\begin{aligned} \Delta F &= \Delta E - T \Delta S \\ &= +2J - k_B T \log(N - 1) \end{aligned}$$

This means that for *any temperature*, if  $N$  is big enough, destroying the ground state energy will lower the free energy! Thus, the entropy will always overwhelm the energetically favoured ordered state. However, note that this argument only works in 1D—the boundary of magnetic domains in 1D is independent of the size of the domain. In  $d \geq 2$ , the boundary grows with the size of the domain, and so there is an additional energy cost which may make spontaneous magnetization favourable.

## 30.4 2D Ising Square Lattice

Rather than derive the exact solution, which is incredibly difficult to understand, we will instead try to approximate the solution. In particular, we will use the *Mean Field Theory* approximation. We will use two MFT approaches—once relying on physical intuition, and again using a powerful formalism.

### 30.4.1 Mean Field Theory—Intuition

Let us assume that we can write the spin on each site as an average

$$s_i = \langle s_i \rangle + \delta s_i = m + \delta s_i$$

where  $m = \langle s_i \rangle$  is the mean magnetization per spin, and  $\delta s_i$  are the fluctuations on site  $i$ . When we plug this into our coupling term  $j$ , we find

$$s_i s_j = m^2 + (\delta s_i + \delta s_j)m + \delta s_i \delta s_j$$

Here, we will assume that fluctuations are uncorrelated. In other words, the expectation of the product of these fluctuations is the product of the expectations:

$$\langle \delta s_i \delta s_j \rangle = \langle \delta s_i \rangle \langle \delta s_j \rangle = 0 \cdot 0$$

However, by definition we chose the  $\delta s_i$  to be centred around zero, and we throw away the final term for simplicity. However, we keep the second term and rewrite our definition so  $s_i - m = \delta s_i$ . Then, our approximation becomes

$$s_i s_j \approx m^2 + m[(s_i - m) + (s_j - m)]$$

or

$$s_i s_j \approx -m^2 + m[s_i + s_j] \tag{30.8}$$

Thus, our hamiltonian becomes

$$H = -J \sum_{\langle ij \rangle} [-m^2 + m(s_i + s_j)] - h \sum_i s_i$$

We note there aren't any more bonds to sum over in the first term. Rather, we use the definition of our lattice to make it a sum over atoms, not bonds. In our lattice, every atom

will have  $z$  neighbours—this is its coordination number—so the lattice will have  $\frac{z}{2}N$  bonds. Thus, we can simplify our first sum by splitting into three sums, and collecting terms, until we obtain

$$H = -\frac{z}{2}NJm^2 - Jzm \sum_i s_i - h \sum_i s_i$$

Thus, our hamiltonian becomes

$$H = \frac{1}{2}NzJm^2 - (h + Jzm) \sum_i s_i \quad (30.9)$$

The first term is a simple shift in the energy, and we will neglect it. On the other hand, we see that the second term acts as an “effective” external field, which is the base field  $h$  with an additional *mean field*  $Jzm$  which comes from the mean configuration of the nearest neighbours. If we break down the mean field, we see that it is the number of neighbours,  $z$ , times their average magnetization  $m$ , times the energy of the coupling,  $J$ .

Thus, we need only study the hamiltonian

$$H = -\tilde{h} \sum_i s_i$$

which we have done before<sup>7</sup>, and has the solution

$$m = \tanh \beta \tilde{h} = \tanh[\beta(h + Jzm)]$$

Note that our solution for the magnetization is a function of itself; this is called a *self-consistency condition*. We want to have the mean magnetization of the spin in question to be the same as the mean magnetization of the bulk.

While this magnetization cannot, in general, be solved for  $m(h)$ , it is relatively straightforward to solve for  $h(m)$  and then reflect the plot over  $m = h$ . Trivially, this function<sup>8</sup> is given

$$h = k_B T \operatorname{artanh}(m) - Jzm \quad (30.10)$$

{diagram}

In particular, it is interesting to consider the case where  $Jzm$  has a slope greater than the slope of  $k_B T \operatorname{artanh}$  and causes the function  $h(m)$  to no longer be injective. This limit is a small temperature limit, while the expected inverse-sigmoid function is a large temperature limit. The temperature where the slopes are exactly equal, or the last temperature at which the function  $h(m)$  is still monotonic is the critical temperature. This occurs when

$$h(m \ll 1) \approx k_B T [m + \mathcal{O}(m^2)] - Jzm$$

$$h'(m = 0) = k_B T - J_z$$

<sup>7</sup>Because we no longer need to consider bonds, this is indistinguishable from the 1D chain with no coupling

<sup>8</sup>Note that the inverse of  $\tanh$  is  $\operatorname{artanh}$ , without a  $c$ . This is because the *ar* stands for *area* rather than what is typically expected for trig functions, *arc*. However, mathematica uses  $\operatorname{arctanh}$ .

so,

$$T_c = \frac{Jz}{k_B} \quad (30.11)$$

Note that the critical temperature is the given by the characteristic temperature of the interaction strength per spin. Further, above the critical temperature, because  $h(m)$  is no longer injective, we cannot invert it.

We instead note that at  $h = 0$ , we have three values of  $m$ ! Just as with the van der Waals gas, the way we find which magnetization state is stable is to choose the state which minimises a thermodynamic potential. Because we have the magnetization from the Helmholtz free energy

$$M = -\frac{\partial F}{\partial h}$$

we can instead perform a Legendre transform to the Gibbs free energy<sup>9</sup>

$$h = +\frac{\partial G}{\partial M}$$

Thus, we can find the free energy as

$$G(m) = \int_0^m dm' h(m')$$

Using this, we can then plot the free energy as a function of the magnetization **{diagram}** We see that for  $T > T_c$  we have a strong quartic behaviour, for  $T = T_c$  we have a flatter quartic, and for  $T < T_c$  we have a mexican hat potential. Then, we see that there are two stable minima, which correspond to  $\pm m_0 \neq 0$ , and we have spontaneous magnetization! Note that the state that corresponds to  $m = 0$  is an unstable state, and thermal fluctuations will cause it to fall into one of the magnetised states.

In another analogy to the vdW gas, we can use the Maxwell equal area construction to cut off the multivalued function part of the graph, and thus we should have at  $h = 0$  the magnetization goes directly from  $-m_0$  to  $+m_0$ , so flipping the sign of the field gives us a phase transition.

---

<sup>9</sup>Note that the gibbs free energy is typically a function of intensive variables, while the helmholtz is a function of extensive variables. Deserno correspondingly swaps the labels to match the intensive and extensive quantities. Here we will not.

### Entropy of the 2D Ising Model

Interestingly the entropy of the Ising model is not necessarily monotonic. To show why we will build a table of the zero-field case. Trivially, the ground state has all aligned bonds, while the highest energy state has all anti-aligned bonds. The highest entropy state is harder to describe. It is easiest to interpret as the “middle” state of a binomial distribution, as there are the largest number of configurations. However, as we have seen with the highest energy state, there are still indeed highly ordered states, but the entropy is dominated by disordered states. Using the sterling approximation,

State	Energy	Degeneracy	Entropy
Ground state	$-JNz/2$	2	$k_B \ln 2$
Highest energy	$JNz/2$	2	$k_B \ln 2$
Highest entropy	0	$\approx \binom{N}{N/2}$	$\approx Nk_B \ln 2$

Clearly, by somewhere between the highest entropy state and the highest energy state the entropy must decrease, and thus we have *negative temperature!* (where we of course recall that  $\frac{1}{T} = \frac{\partial S}{\partial E}$ ). The analytic form of this function as  $N \rightarrow \infty$  is given

$$\frac{s}{k_B} \sim \frac{e}{J} \ln \frac{e}{J}$$

We see that when we take the derivative, above  $e/J = 0$ , we have only negative temperatures. We see that the lowest temperature that the Ising model can have is  $T = 0^+$  and the lowest  $T = 0^-$  where the sign denotes the direction the temperature approaches absolute zero; it is often, then, usefull to consider  $\beta$  as the temperature instead, as it is continuous for all energies, in that it doesn't go from  $+\infty$  to  $-\infty$  at  $e/J = 0$ .

### 30.4.2 Mean Field Theory—Variational Principles

Recall the result we showed on Exam 1 using the gibbs inequality:

$$F \leq F_t + \langle H - H_t \rangle_t \quad (30.12)$$

Or, the true free energy is bounded above by an expression derived using a trial Hamiltonian  $H_t$ . If we parametrize  $H_t$ , we can then make it into a variational principle. Further, although in general taking canonical averages is difficult, *we can choose the trial hamiltonian to make it easy*. Namely, we choose a trial hamiltonian without interactions.

For the Ising model, we take the trial hamiltonian to be

$$H_t = -\alpha \sum_i s_i$$

the solution is the same as for the no-coupling ising model:

$$Z_t = (2 \cosh \beta \alpha)^N$$

$$F_t = -Nk_B T \log(2 \cosh \beta \alpha)$$

$$m = \tanh \beta \alpha$$

Now we do the computation involving the true hamiltonian

$$H = - \sum_{\langle ij \rangle} s_i s_j - h \sum_i s_i$$

we compute

$$\langle H - H_t \rangle_t = \sum_{\{s\}} \left( -J \sum_{\langle ij \rangle} s_i s_j - (h - \alpha) \sum_i s_i \right) \frac{e^{\beta \alpha \sum_i s_i}}{Z_t} \quad (30.13)$$

Because the trial state factorizes, we can similarly factorize the expectation value<sup>10</sup>:

$$= -J \sum_{\langle ij \rangle} \langle s_i \rangle_t \langle s_j \rangle_t - (h - \alpha) \sum_i \langle s_i \rangle_t \quad (30.14)$$

$$= -\frac{1}{2} N z J \langle s \rangle_t^2 - (h - \alpha) N \langle s \rangle_t \quad (30.15)$$

Recognizing that  $\langle s \rangle_t = m$  in the trial state,

$$= -\frac{1}{2} N z J \tanh^2 \beta \alpha - (h - \alpha) N \tanh \beta \alpha \quad (30.16)$$

Thus, our overall variational principle is given

$$F \leq -Nk_B T \log(2 \cosh \beta \alpha) - \frac{1}{2} N z J \tanh^2 \beta \alpha - (h - \alpha) N \tanh \beta \alpha$$

Minimising the RHS wrt  $\alpha$ ,

$$\begin{aligned} 0 &= \frac{\partial}{\partial \alpha} (F_t + \langle H - H_t \rangle_t) \\ &= \cancel{-Nk_B T \frac{2 \sinh \beta \alpha}{2 \cosh \beta \alpha}} \beta - N z J \frac{\tanh \beta \alpha}{\cosh^2 \beta \alpha} \beta + \cancel{N \tanh \beta \alpha} - (h - \alpha) N \frac{\beta}{\cosh^2 \beta \alpha} \\ &= -N z J \frac{\tanh \beta \alpha}{\cosh^2 \beta \alpha} \beta - (h - \alpha) N \frac{\beta}{\cosh^2 \beta \alpha} \\ &= -J z \tanh \beta \alpha - (h - \alpha) \end{aligned}$$

so

$$\frac{h - \alpha}{J z} = \tanh \beta \alpha$$

Note again that the RHS is simply  $m$ . Thus, the LHS as  $m$  lets us write  $\alpha = h + J z m$  and we reobtain our self-consistency relation

$$m = \tanh[\beta(h + J z m)]$$



# Chapter 31

## Critical Phenomena

### 31.0.1 Validity of MFT

Note that MFT predicts a phase transition for  $z = 2$ . However, we know the *exact* answer for  $z = 2$  as the 1D Ising chain, *which has no transition!* Thus, MFT is *qualitatively wrong* in 1D. So then, can we trust the results of MFT for  $d > 1$ ? Indeed, the existence of the phase transition in 2D was proven by Robert Griffiths, who fixed (pyle’s?) argument. Then, people used the low temperature series expansion, and similar high temperature series expansion, and drew correlations between the two to extrapolate the phase transition, by showing the two series begin to diverge at the same point. This gave validity to MFT as a technique. However, Onsager’s exact solution to the 2D lattice was concerning to MFT. Further rigorous validations were conducted using series expansion, such as what was done in HW. Such series are exact to the point they are truncated.

### 31.1 Critical Exponents

The behaviour of thermodynamics in the vicinity of a critical point (such as a 2nd order phase transition) is often described, to leading order by power laws. Such a critical point is where the two phase coexistence regions contract to a point (such as the two branches of the Ising model in the bond approximation). For example, in fluids, the critical point is the boundary between a two-phase coexistence region and a continuous transition through a supercritical fluid. The exponent of these power laws are known as *critical exponents*.

These critical exponents are thus extremely important in describing thermodynamic quantities. Most interestingly, however, these critical exponents display a “universality” in that they depend solely on the dimension of space  $d$ , and the units of the *order parameter*. The order parameter for a vdW fluid, for instance, is density; as we approach the critical point, this order parameter vanishes. Similarly, for magnetic states, we can use the magnetization as the order parameter, and as we approach the critical point, the difference in the magnetization between the two states vanishes; for alloy systems, we can use mean occupation by species as the order parameter. Yet, for all these systems, *the critical exponent is the same*

(for fixed  $d$ ). These critical exponents are denumerated by greek letters.

## Heat Capacity

The first exponent used is  $\alpha$ ; this is used for the leading order in heat capacity:

$$c \sim \left| \frac{T - T_c}{T_c} \right|^{-\alpha} \equiv |t|^{-\alpha} \quad (31.1)$$

Where we introduce

$$t \equiv \frac{T - T_c}{T_c} \quad (31.2)$$

Recall that from the fluctuation response theorem,

$$c \propto \sigma_E^2$$

Thus, when  $c$  diverges (namely at  $T_c$ ), the energy fluctuations similarly diverge! This breaks two things: First, in MFT, where we assume that these fluctuations are separable, our assumptions break down. Second, we can no longer apply Central Limit Theorem, as CLT assumes that the variance is finite<sup>1</sup>

Recall that

$$c = \left( \frac{\partial E}{\partial T} \right)$$

so

$$E(T) = \int^T dT' c(T')$$

which means there is a boundary on how much the heat capacity can diverge. In particular,  $\alpha < 1$ . Further, we see that  $E(T_c)$  must be vertical. For example, recall the phase transition ice→water. At 0°C we can pump energy in without changing the temperature—all of the energy goes into *latent heat* which is used to drive the phase transition, rather than increasing the temperature of the system. For a first order transition, this region is finite, but for a continuous transition, this occurs at a point. In particular, for a first order phase transition, this region is satisfied by adding a delta function at the transition temperature. In fact, the heat capacity doesn't even need to diverge (beside the delta function).

## Zero-field Magnetization

The critical exponent  $\beta$  determines how the magnetization varies with temperature at zero field strength.

$$m \sim (-t)^\beta \quad (31.3)$$

---

<sup>1</sup>The Central Limit theorem states that for a random variable  $x$ , and a second variable  $X = \sum_{i=1}^N x_i$ , as  $N \rightarrow \infty$  the probability distribution  $P_X$  approaches a gaussian, with variance the sum of the individual variances. This applies no matter what  $P_x$  is.

Of course, for  $T > T_c$  we have no spontaneous magnetization, but below, this is where we have our critical exponent dependence.

This is very similar qualitatively to the macroscopic occupation of a Bose-Einstein condensate. In particular, when we expand about  $T \sim T_E$ , the BEC has an exponent  $\beta = 1$ . For a magnetic system, this order parameter is typically  $< 1$ , but can in principle be  $= 1$ . This is the behaviour of a single branch of the Ising model.

### Zero field isothermal susceptibility

The critical exponent  $\gamma$  determines how the magnetic susceptibility changes at zero field:

$$\chi_T = \left( \frac{\partial m}{\partial h} \right)_T \quad h = 0 \quad (31.4)$$

This is the *response* of the magnetization to the driving of the field.

$$\chi_T \sim |t|^{-\gamma} \quad (31.5)$$

We once again obtain a divergence at  $T = T_c$ , which corresponds to the divergence of the associated fluctuations  $\sigma_m$ . In fact, this is *exactly* what we observed when we plotted  $h(m)$ ; there is a discontinuous jump between the two spontaneous magnetization states.

### Critical Isotherm

The exponent  $\delta$  is defined via

$$h(m) \sim \text{sgn}(m) |m|^\delta \quad (31.6)$$

Note that this is related to  $\gamma$  in definition, but fixes  $T$  rather than  $h$ .

### Spatial Correlations

The last two exponents are related to spatial correlations. In particular, consider two separated spins,  $\sigma_i$  and  $\sigma_j$ , separated by  $r_{ij} = r_j - r_i$ . We can then define a correlation function

$$\Gamma(\mathbf{r}_{ij}) = \langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle \quad (31.7)$$

which is the covariance of the two spins. If they are uncorrelated, we expect  $\Gamma = 0$ . If the spins are sufficiently far apart, we should expect  $\Gamma(r \rightarrow \infty) \rightarrow 0$ . Although nearby spins are strongly correlated, as these correlations “telephone” the strength of the correlations decreases over time. We are concerned with the *correlation length* after which the spins are essentially uncorrelated.

Thus, (for most materials)

$$\Gamma(\mathbf{r}) \rightarrow \Gamma(r) = \frac{e^{-r/\xi}}{r^\tau} \quad (31.8)$$

the length  $\xi$  is called the *correlation length*, and describes how quickly the correlations die out.

As temperature decreases, entropic effects die out, and we see that the correlation length should in general increase. At the critical temperature, the length  $\xi$  should diverge, and  $\Gamma \rightarrow r^{-\tau}$ . Below the critical temperature, the correlations once again decrease. This is somewhat non-intuitive.

The critical exponent  $\nu$  determines the correlation length via

$$\xi \sim |t|^{-\nu} \quad (31.9)$$

If the correlation length diverges at  $T = T_c$ , we have extremely long range “memory” of correlations in our systems which decays as a power law rather than an exponential. However, at the critical temperature, the exponent  $\tau$  changes—rather, we have our second critical exponent which describes our correlation decay at the critical temperature:

$$\Gamma(r) \sim \frac{1}{r^{d-2+\eta}} \quad (31.10)$$

or,  $\eta$  is the correction to the “simplistic” expectation.

## 31.2 Landau Theory

Landau theory, while its name might be intimidating, is an extremely elegant, simple, and quick way to build a mean field theory. To do this, we expand the free energy in powers of the order parameter, accounting for symmetries. Recall that an order parameter is a thermodynamic observable that distinguishes two phases; for example, density distinguishes the phases of a vdW fluid, and magnetization can distinguish paramagnets from ferromagnets. In particular, order parameters discontinuously jump across a first order phase transition, or has a vertical tangent for continuous phase transitions.

### 31.2.1 Ferromagnet

The order parameter of a ferromagnet is the magnetization. In particular, let us first consider the zero field case. Then, the symmetry of the system invariance regardless of upward orientation or downward orientation. Thus, the sign of  $m$  doesn’t matter, and we know that our series expansion must only hold even powers of  $m$ . Thus, we take the ansatz

$$F = F_0 + a_2 m^2 + a_4 m^4 + \dots \quad (31.11)$$

In particular, our hamiltonian is invariant under a transformation  $\sigma_i \rightarrow -\sigma_i$

$$H = J \sum_{\langle ij \rangle} \sigma_i \sigma_j \rightarrow J \sum_{\langle ij \rangle} (-\sigma_i \sigma_j) = J \sum_{\langle ij \rangle} \sigma_i \sigma_j = H$$

If  $a_2 > 0$ , a negative quartic term causes the free energy to be unbounded below, and thus we must have  $a_4 \geq 0$  and so the free energy is minimised at  $\langle m \rangle = 0$

If instead  $a_2 < 0$  to keep the free energy bounded below, we must have  $a_4 > 0$ , which gives us a mexican hat potential, and spontaneous magnetization, or a ferromagnetic phase.

If we vary the system continuously between these two states, we find that the phase transition occurs at  $a_2 = 0$ . In particular, if we consider

$$a_2 = \tilde{a}_2 \cdot t \quad (31.12)$$

with  $\tilde{a}_2 > 0$  and  $t = (T - T_c)/T_c$ , we obtain a very powerful theory of phase transitions. In fact, we can predict basically all critical exponents!

First, we find the magnetization. If  $t > 0$ , we trivially have  $m = 0$ . For  $t < 0$ , we minimise the free energy

$$0 = \frac{\partial F}{\partial m} = 2\tilde{a}_2 t m + 4a_4 m^3$$

so

$$m = \pm \sqrt{-\frac{\tilde{a}_2}{2a_4} t} \quad t < 0$$

Thus,

$$m \sim (-t)^\beta = (-t)^{1/2} \quad (31.13)$$

or  $\boxed{\beta = 1/2}$  Inserting this into the free energy, we have

$$F = F_0 - \frac{\tilde{a}_2^2}{2a_4} t^2 + a_4 \frac{\tilde{a}_2^2 t^2}{2a_4} = F_0 - \frac{\tilde{a}_2^2 t^2}{4a_4}$$

We can find the heat capacity by

$$c = \frac{T}{N} \frac{\partial S}{\partial T} = -\frac{T}{N} \frac{\partial^2 F}{\partial T^2}$$

so

$$\lim_{t \rightarrow 0^-} c = \frac{\tilde{a}_2^2}{2NT_c a_4} \quad (31.14a)$$

however, for  $t > 0$ ,  $m = 0$  and so  $F = F_0$  and

$$\lim_{t \rightarrow 0^+} c = 0 \quad (31.14b)$$

Thus, the specific heat has a discontinuity at  $T = T_c$  which corresponds to

$$\alpha = 0 \quad (31.14c)$$

This comes from a more formal definition the critical exponent in a function  $f(t)$  as

$$\alpha = \left. \frac{\partial \log[f(t)]}{\partial \log t} \right|_{t=t_0}$$

Note however, a logarithm  $\log(t)$  can also correspond to  $\boxed{\alpha = 0}$ .

Now, if we apply a field, our free energy gains a linear term:

$$F \approx F_0 - hm + \tilde{a}_2 tm^2 + a_4 m^4 \quad (31.15)$$

Of course, we find equilibrium by minimizing  $F$  wrt  $m$ :

$$0 = \frac{\partial F}{\partial m} = -h + 2\tilde{a}_2 tm + 4a_4 m^3$$

In particular, on the critical isotherm,  $T = T_c \implies t = 0$ , we have

$$h = 4a_4 m^3$$

and so we find that  $\boxed{\delta = 3}$ .

When  $t \neq 0$ , we have

$$h = 2\tilde{a}_2 tm + 4a_4 m^3$$

and so we find our susceptibility

$$\begin{aligned} \frac{\partial h}{\partial m} &= 2\tilde{a}_2 t + 12a_4 m^2 \\ \chi &= \frac{\partial m}{\partial h} = \frac{1}{2\tilde{a}_2 t + 12a_4 m^2} \end{aligned}$$

when  $t > 0$ , we know that  $m = 0$ , so

$$\chi_+(t) = \frac{1}{2\tilde{a}_2 t} \sim t^{-1}$$

however, for  $t < 0$ , our zero-field magnetization is  $m^2 = -\frac{\tilde{a}_2 t}{2a_4}$ , so

$$\chi_-(t) = -\frac{1}{4\tilde{a}_2 t} \sim t^{-1}$$

Note that the negative sign cancels with the negative temperature, and so the susceptibility is still positive. In both cases, we see that  $\boxed{\gamma = 1}$ . It is interesting, however, that the prefactors are different on either side of  $t = 0$ . In particular,

$$\frac{\chi(t \rightarrow 0^+)}{\chi(t \rightarrow 0^-)} = 2$$

Thus, while the critical exponents are universal, the prefactors are in general not.

Up to this point, we have not had to consider the spatial dependence of our lattice; we have not needed to care about the number of dimensions in the model. However, to calculate the exponents  $\nu$  and  $\eta$ , we now need to make our Landau model *local*. Thus, we define a “free energy density”  $f$  and define the free energy density dependent on a magnetization  $m(\mathbf{r})$  and integrate to obtain the full free energy. To penalize a rapidly changing magnetization,

we add a term that is the squared modulus of the gradient. This is known as a *Ginzburg Landau Theory*<sup>2</sup>. Our free energy can be written

$$F - F_0 = \int d^d r [g(\nabla m)^2 + \tilde{a}_2 t m^2 + a_4 m^4] \quad (31.16)$$

Let us consider first  $t > 0$ , where we don't care about  $a_4 m^4$ . We can then go to Fourier space, as we know  $e^{i\mathbf{kr}}$  are eigenectors of the gradient operator<sup>3</sup>. Recall that we have

$$\int d^d r e^{-i\mathbf{q}\cdot\mathbf{r}} = (2\pi)^d \delta(\mathbf{q})$$

Our magnetization is transformed

$$m(\mathbf{r}) = \frac{1}{(2\pi)^d} \int d^d q e^{i\mathbf{q}\cdot\mathbf{r}} \tilde{m}(\mathbf{q})$$

$$m(\mathbf{r}) = \int d^d r e^{-i\mathbf{q}\cdot\mathbf{r}} m(\mathbf{r})$$

Plugging into our free energy, we find

$$F - F_0 = \int \frac{d^d q}{(2\pi)^d} (\tilde{a}_2 t + dq^2) |\tilde{m}(\mathbf{q})|^2 \quad (31.17)$$

This gives us a “sum” of independent quadratic terms  $|\tilde{m}|^2$ , with an effective spring constant. Using the equipartition theorem, we see for each degree of freedom

$$\langle (\tilde{a}_2 t + gq^2) |\tilde{m}(\mathbf{q})|^2 \rangle = k_B T$$

Up to a factor<sup>4</sup> 1/2. From this result, we are able to find the correlation function

$$\Gamma = \langle m(\mathbf{r})m(0) \rangle - \langle m(\mathbf{r}) \rangle \langle m(0) \rangle \rightarrow \langle m(\mathbf{r})m(0) \rangle$$

where we are able to throw out the second term because at  $t > 0$  we have  $\langle m \rangle = 0$ . Considering the fourier transform,

$$\begin{aligned} \tilde{\Gamma}(\mathbf{q}) &= \langle \tilde{m}(\mathbf{q})m(0) \rangle \\ &= \int \frac{d^d q'}{(2\pi)^d} \langle \tilde{m}(\mathbf{q})\tilde{m}(\mathbf{q}') \rangle \end{aligned}$$

However, as individual modes  $m(\mathbf{q})$  are independent, we can rewrite the expectation<sup>5</sup> as

$$= \int \frac{d^d q'}{(2\pi)^d} \delta(q + q') (2\pi)^d \langle |\tilde{m}(\mathbf{q})|^2 \rangle$$

<sup>2</sup>Such models are used to study superconductive phases

<sup>3</sup>Unfortunately, the source of this derivation used a different convention from Deserno. Luckily, these are all multiplicative prefactors, and do not affect the value of the critical exponents.

<sup>4</sup>Deserno was unable to figure out where the factor went; perhaps the two degrees come from complex  $\tilde{m}$ ?

$$\begin{aligned}
&= \langle |\tilde{m}(\mathbf{q})|^2 \rangle \\
&= \frac{k_B T}{\tilde{a}_2 t + g q^2}
\end{aligned}$$

To return to real space, we can take an inverse Fourier transform. The asymptotic behaviour goes with

$$\Gamma(\mathbf{r}) \sim \begin{cases} \frac{e^{-r/\xi}}{r^{(d-3)/2}} & t > 0 \\ \frac{1}{r^{d-2}} & t = 0 \end{cases} \quad (31.18)$$

where

$$\xi = \sqrt{\frac{g}{\tilde{a}_2 t}}$$

Thus, we find the two critical exponents  $\boxed{\nu = 1/2}$  and  $\boxed{\eta = 0}$ .

In summary, the critical exponents we have calculated using Landau theory are

$$\alpha = 0 \quad (31.19a) \quad \beta = 1/2 \quad (31.19c) \quad \gamma = 1 \quad (31.19e)$$

$$\delta = 3 \quad (31.19b) \quad \nu = 1/2 \quad (31.19d) \quad \eta = 0 \quad (31.19f)$$

### 31.3 MFT Validity

We began this discussion of critical exponents with the question of “When is MFT valid?” In particular, MFT ignores fluctuations, and so we expect it to be valid only when the fluctuations are unimportant. However, near critical points, we see that fluctuations actually diverge.

In particular, at a critical point, we can write the free energy density of fluctuations as

$$f_{fluc} \sim \frac{k_B T}{\xi^d} \sim |t|^{\nu d}$$

where  $k_B T$  gives us the energy, and the correlation length  $\xi$  gives us a characteristic length scale.

In addition, we can integrate the specific heat twice to compute the actual free energy density

$$f \sim |t|^{2-\alpha}$$

In particular, for MFT, we want for the fluctuations to be small compared to the free energy:

$$1 \ll \frac{f_{fluc}}{f} \sim \frac{|t|^{\nu d}}{|t|^{2-\alpha}} \sim |t|^{\nu d - 2 + \alpha}$$

At the critical point,  $t \rightarrow 0$ , and so MFT is valid iff

$$0 < \nu d - 2 + \alpha$$



so we obtain a self-consistency requirement

$$\nu d > 2 - \alpha \quad (31.20)$$

We can plug in the values we obtained from Landau theory to find that

$$\frac{d}{2} > 2$$

which means that MFT of the Ising model is only valid when  $d > 4$ , even *at* the critical temperature. In fact, for  $d = 4$ , it is correct up to logarithmic corrections. This is known as the *upper critical dimension*, the minimum dimension for which MFT is valid. There are even some systems where this upper critical dimension is 3 and we can use MFT on real-world systems.

## 31.4 Critical Exponent Equalities

It turns out the critical exponents are not mutually independent; rather, they satisfy a number of identities. The first of these is

$$\alpha + 2\beta + \gamma = 2 \quad (31.21)$$

which is known as the Rushbrooke's identity. It is rather simple to prove as an inequality due to convexity consideration, but it is interesting that it is actually an equality. Another is

$$\beta(\delta - 1) = \gamma \quad (31.22)$$

This is known as *Widom's Identity*, after Prof. Mike Widom's father, Ben Widom. The (in)equality we just used to show the consistency of MFT

$$2 - \alpha = d\nu \quad (31.23)$$

is known as Josephson's identity; in fact, one of the reasons MFT is questionable is because it violates Josephson's identity. Finally, we will list one more identity, whose name is unknown

$$(2 - \eta)\nu = \gamma \quad (31.24)$$

The amazing insight that Ben Widom (and indeed was well known for) was that these relations follow naturally if the singular part of the free energy near the critical temperature follows a scaling relation. For the Ising model, this scaling relation is given

$$f_s(t, h) = t^{2-\alpha} g_f \left( \frac{h}{t^\Delta} \right) \quad (31.25)$$

where  $\Delta$  is known as the *gap exponent*. This incredibly restrictive form of the leading order singular behaviour of the free energy gives us these equalities<sup>6</sup> The energy is found

$$E_s \sim \frac{\partial f_s}{\partial t} = (2 - \alpha)t^{1-\alpha} g_f \left( \frac{h}{t^\Delta} \right) - \Delta h t^{1-\alpha-\Delta} g'_f \left( \frac{h}{t^\Delta} \right)$$

---

<sup>6</sup>We sometimes say that “ $h$  enters only in the combination  $h/t^\Delta$ .”

Factoring out a factor of  $t^{1-\alpha}$ ,

$$= t^{1-\alpha} \left[ (2-\alpha)g_f\left(\frac{h}{t^\Delta}\right) - \Delta \frac{h}{t^\Delta} g'_f\left(\frac{h}{t^\Delta}\right) \right]$$

Thus, we can introduce a new scaling function  $g_E$ :

$$= t^{1-\alpha} g_E\left(\frac{h}{t^\Delta}\right)$$

Similarly, we find the dependence of the heat capacity as

$$c_s \sim t^{-\alpha} g_c\left(\frac{h}{t^\Delta}\right) \quad (31.26)$$

As  $h \rightarrow 0$ , we see that  $g_c$  goes to some constant, and we recover our critical exponent  $\alpha$ . By construction, this is true, and not very interesting. However, we can find the other exponents in a similar manner. For instance, the magnetization can be found

$$m_s(t, h) = \frac{\partial f_s}{\partial h} \sim |t|^{2-\alpha-\Delta} g_m\left(\frac{h}{t^\Delta}\right) \quad (31.27)$$

and with  $h \rightarrow 0$ , we obtain our second critical exponent

$$\beta = 2 - \alpha - \Delta \quad (31.28)$$

Similarly, as  $t \rightarrow 0$  we find  $h/t^\Delta \rightarrow \infty$ . We will assume that  $g(x) \sim x^p$  as  $x \rightarrow \infty$ . Thus,

$$m(t=0, h) \sim |t|^{2-\alpha-\Delta} \frac{h^p}{t^{\Delta p}} = h^p |t|^{2-\alpha-\Delta-p\Delta}$$

However, the LHS is independent of  $t$ , while the RHS is dependent on  $t$ . Thus, we conclude the exponent is zero, and we must have

$$p\Delta = 2 - \alpha - \Delta \quad (31.29)$$

Thus,

$$m_s(t=0, h) \sim h^p = h^{2-\alpha-\Delta/\Delta} = h^{\beta/\Delta}$$

inverting, we find

$$h = [m_s(t=0, h)]^{\Delta/\beta}$$

so

$$\delta = \frac{\Delta}{\beta} \quad (31.30)$$

Next, we find the susceptibility

$$\chi_s(t, h) \sim \frac{\partial m_s}{\partial h} \sim |t|^{2-\alpha-2\Delta} g_\xi\left(\frac{h}{t^\Delta}\right)$$

The zero field susceptibility scales with

$$\xi_s(t, h=0) \sim |t|^{2-\alpha-2\Delta}$$

so

$$\gamma = 2\Delta - 2 + \alpha \quad (31.31)$$

Using these equations, we can eliminate the  $\Delta$ s and obtain the identities discussed earlier. For example, we note that twice 31.28 gives us Rushbrooke's identity, and subtracting one from 31.30. In particular, these results are from a new “toy” that theorists had at the time—these identities are a triumph of *renormalization theory*.

## 31.5 Actual Values

We will conclude our discussion of critical phenomena by comparing the values of the critical exponents.

		$\alpha$	$\beta$	$\gamma$	$\delta$	$\nu$	$\eta$
MFT		0(jump)	1/2	1	3	1/2	0
2D Ising	2-comp scalar	0(logarithm)	1/8	7/4	15	1	1/4
3D ising	2-comp scalar	0.1	0.33	1.24	4.8	0.63	0.04
2D Potts	q=3 comp scalar	1/3	1/5	13/9	13	5/6	4/15
2D Potts	q=4 comp scalar	2/3	1/12	7/6	15	2/3	1/4

We see that the exponents obtained by exact solutions varied from MFT by a *huge* amount. This led to a lot of work on explaining the failures of MFT, such as new series expansions, to improve these approximation methods.