

WHAT HAPPENED AFTER WE DISCOVERED FIRE

AN OVERVIEW OF THERMODYNAMICS

PHY 301 – THERMAL PHYSICS
RAVIN RAJ

Contents

0	Prelude	2
1	The First Law of Thermodynamics	4
2	The Second Law of Thermodynamics	9
3	Thermodynamic Potentials	26
4	The Third Law of Thermodynamics	35
5	Open Systems	38
6	Phases	43
	References	59

Preface

Please note that this set of notes is **NOT** meant to be a substitute for any course material, and is instead complementary to it at best. The following review is mostly a transcription of some handwritten notes I kept as an undergraduate when I took my first and second statistical mechanics courses. It is hard for me to place exact citations throughout this, as I do not include those in my own notes. As far as possible, I have tried to compile a list of the references [1–9] which I used when writing these and taking my statistical mechanics course in the past. This document may be updated periodically, and the latest version will be put on Canvas as well as my [personal webpage](#).

In my personal opinion, the history and study of thermodynamics is incredibly beautiful and profound. All energy is created equal, yet the idea that some energy (heat) is less equal than others (work) is rather baffling. We take this simple fact for granted in our everyday lives, and the formal study of thermodynamics ought to open your eyes to this. In this course, you would have seen the idea of entropy from a statistical mechanics perspective and how it links to the microscopic constituents of a system. Thermodynamics is a field conceived far before we were even aware of the existence of the microscopic world, and so it approaches heat from a radically different angle. In studying thermal physics as a whole, one should be aware of these two narratives and how they ultimately tell the same story under a different set of assumptions. The ideas developed here tell us how energy and entropy ultimately determine the fate of systems and processes, and one can apply these notions to understand far more complex fields such as biological physics, cosmology, information theory and even sociology.

As mentioned in the previous review, my first (and second and third) statistical mechanics course(s) brought me plenty of joy, in that it/they really did change the way I viewed the world around me. My aim with this review is to (hopefully) pass some of that excitement on to you.

0 Prelude

Before we begin, allow me to set the stage with some preliminary knowledge that I believe you should have going into this. Feel free to skip this section if you feel confident with this content.

0.1 Classes of Thermodynamic Systems

A **thermodynamic system** is an arbitrary amount of matter confined by walls of some sort against its surroundings. **Isolated** systems are those which have walls that prevent any form of interaction with their surroundings, be it in the form of heat exchange, volume exchange, particle exchange or information exchange. Thus, energy is conserved over time in an isolated system. In reality, the closest thing to an isolated system is the entire Universe as a whole (as far as we know). While one may think a perfect vacuum devoid of particles is a great example of an isolated system, even a vacuum can induce spontaneous fluctuations in the underlying quantum fields of the Universe, giving rise to pair production and annihilation, and is also always filled with blackbody radiation at any nonzero temperature. Thus, the perfect isolated system is merely just an idealisation and should not be taken too seriously beyond theoretical work.

Closed systems, on the other hand, are allowed to only exchange energy with their surroundings through heat, so they do not obey conservation of energy. If a closed system is in equilibrium¹ with its surroundings, the mean value of its internal energy will be related to its temperature, or the temperature of the surroundings (which typically sets that of the system).

Finally, **open systems** can exchange both energy and matter with their surroundings, so both the energy and particle number are not conserved². If an open system is in equilibrium with its surroundings, the mean values of the internal energy and particle number will be related to the temperature and chemical potential of the surroundings respectively.

In order to proceed with thermodynamics proper, we need a working definition of thermodynamic equilibrium (we will build on this later). The **thermodynamic equilibrium** state is defined to be the one macroscopic state of the system that is *automatically* attained after a sufficiently long period of time such that the macroscopic properties of the system no longer change over time. For this to occur, we need to specifically break down thermodynamic equilibrium into three distinct requirements:

Definition 0.1: Constituents of Thermodynamic Equilibrium

1. **Thermal equilibrium** is achieved when the temperature differential between the system and its surroundings is zero and unchanging over time.
2. **Mechanical equilibrium** is achieved when there are no unbalanced forces acting on any part of the system or on the system as a whole.
3. **Diffusive equilibrium** is achieved when there are no chemical reactions within the system and there is no net motion of particles from one part of the system to another.

¹We will formalise the definition of equilibrium later on.

²In reality, every system that we can study in the lab is an open system. While one may say that state-of-the-art ultrahigh vacuum chambers are able to prevent the exchange of particles between the interior and exterior of the chamber, even the highest grades of stainless steel inevitably allow for hydrogen to slowly diffuse through the metal into the chamber. This takes place on the time scale of several years, so you have a great approximation to a closed system for awhile before you have to worry about particle exchange.

0.2 State Variables

State variables, or **state functions**, are macroscopic quantities which characterise a system. They are only well defined when the system is in equilibrium, and are only dependent on the current thermodynamic state of the system (with no reference to its past history). The **equation of state** for a system gives a mathematical relationship between state variables³. State variables are typically classified by their degree of homogeneity, where we define a **homogeneous function** as:

Definition 0.2: Homogenous Function

A function $f = f(\{x_i\})$ is said to be a homogeneous function of degree k if it satisfies:

$$f(\{\lambda x_i\}) = \lambda^k f(\{x_i\}), \quad \lambda \in \mathbb{R} \quad (0.1)$$

Homogeneous functions of degree 1 are called **extensive functions**, while those of degree 0 are called **intensive functions**.

In the context of thermodynamics, we note that extensive state variables are proportional to the amount of matter in a system, while intensive state variables are independent of the amount of matter. Finally, we also note that all state variables in thermodynamics must produce **exact differentials**, since the presence of an inexact differential would imply that the variable depends on the path taken in some state space.

0.3 The Zeroth Law of Thermodynamics

The zeroth law of thermodynamics, just by virtue of its name, sounds like a hilarious afterthought (and it probably is, granted its importance). It reads as follows:

Definition 0.3: Zeroth Law of Thermodynamics

If two systems are separately in thermal contact and equilibrium with a third system, then they must also be in thermal equilibrium with one another.

This implies the existence of a state variable which encodes the propensity for two systems in thermal contact to exchange energy with one another — the **temperature** T . That is, two systems with different initial temperatures will spontaneously exchange energy with one another once brought into thermal contact, until thermal equilibrium is achieved. As history would have it, there are various temperature scales in use around the world such as the Kelvin scale, the Celsius scale, the Fahrenheit scale (and my personal favourite, the Newton scale⁴). We will exclusively work in terms of degrees Kelvin, which has a fixed *calibration point* defined by the **triple point** of water (we will formalise the definition of the triple point later) at 273.16 K. The Kelvin scale for an ideal gas is then defined as:

$$T = \frac{\lim_{P \rightarrow 0} (PV)_{\text{system}}}{\lim_{P \rightarrow 0} (PV)_{\text{ice-water-steam}}} \times 273.16 \text{ K} \quad (0.2)$$

³Equations of state are typically given empirically in thermodynamics, while statistical mechanics provides us with a bottom-up approach for constructing these equations of state from the microscopic constituents of a system and their governing equations of motion.

⁴In the Newton scale, 0 is defined as “the heat of air in winter at which water begins to freeze”, while 6 is defined as “the heat at midday about the month of July”.

1 The First Law of Thermodynamics

Now that that's out of the way, let's begin.

1.1 Internal Energy

We recall from classical mechanics that the energy of a system can be broken into:

$$E_{\text{total}} = E_{\text{macroscopic}} + E_{\text{microscopic}} \quad (1.1)$$

The macroscopic energy comprises the kinetic energy of the centre of mass and the potential energy of the system due to external fields only. In contrast, the microscopic energy comprises all forms of energy resulting from the motion of the microscopic constituents of a body or system. As an example, the prototype problems in elementary mechanics typically make reference to a solid block undergoing some sort of motion due to external forces. The macroscopic energy of this block would simply be the sum of the kinetic and potential energies of the block's centre of mass, while the microscopic energy would be that due to the vibrations and interatomic repulsion of the individual atoms making up the block. While in classical mechanics, we disregard the microscopic degrees of freedom to study macroscopic bodies, we will do the opposite in thermodynamics and instead disregard the macroscopic degrees of freedom such as motion and interactions of the centre of mass. Thus, it is automatically assumed that the energy being referenced in thermodynamics and statistical mechanics is purely microscopic in origin, so we identify $E = E_{\text{microscopic}}$.

The internal energy of a system is the sum of the energy of all internal degrees of freedom of the system. We can break this up into the kinetic energy E_{kin} and the potential energy E_{pot} . The kinetic energy accounts for all microscopic constituents' kinetic energies in a frame where the centre of mass is at rest, while the potential energy accounts for the interactions between all constituent particles in the system. Now, we also note that the internal energy E is an extensive state function and depends only on the equilibrium state of the system described by some set of functions $\{X_i\} = \mathbf{X}$ (represented by a coordinate vector in state space). We then have $E = E(\mathbf{X})$, where \mathbf{X} is a full characterisation of the system in the relevant state space, independent of the system's history or previous path through the state space.

1.2 The First Law

In order to proceed, we need to first recognise that microscopic energy primarily comes in two forms. Specifically, if energy flows between a system and its surroundings due to a temperature differential across a wall, we consider this to be **heat** Q . All other forms of energy transfer are then considered to be **work** W . We will later see that heat is solely responsible for changing the amount of disorder in a system, while work has no effect on the disorder.

In an ideal gas, the natural state space to consider is that of the pressure p , volume V and temperature T . For now, we will simply consider the two-dimensional subspace parametrised by (p, V) . Each point in a pV -diagram is described by the two coordinates p and V , since they are state variables. However, both the heat and work are not state variables, and depend on the nature of the underlying process and the path that they take through this state space. This motivates the first law of thermodynamics, which relates the path variables of heat and work, to the state variable of internal energy.

Definition 1.1: First Law of Thermodynamics

The **first law of thermodynamics** states that energy is conserved, and is transferred between systems as either heat or work. The increase in internal energy of a system is thus equal to the sum of heat flowing into it and the work done on it. Mathematically, we have:

$$\delta E = E(\mathbf{X}_f) - E(\mathbf{X}_i) = \delta Q + \delta W \quad (1.2)$$

where δ represents a finite change in a quantity. In differential (infinitesimal) form, we equivalently write:

$$dE = \mathring{d}Q + \mathring{d}W \quad (1.3)$$

where the use of \mathring{d} implies that the differential is for a path variable (rather than a state variable), and is thus not exact. We will also use the sign convention that Q and W (and their changes δ or differentials \mathring{d}) are positive for heat flowing into a system or work being done on the system.

The fact that dE is an exact differential implies that the inexact differentials of heat and work must be expressible in terms of suitable state variables, specifically of the form:

$$dE = \sum_i \left(\frac{\partial E}{\partial X_i} \right)_{\{X_j\}_{j \neq i}} dX_i \quad (1.4)$$

where the subscript on a partial derivative lists the variables to be kept constant in taking the derivative.

1.3 Quasistatic and Reversible Processes

Definition 1.2: Quasistatic and Reversible Processes

A process is said to be **reversible** if it is possible to restore the system and its surroundings to their original conditions (note the emphasis on the surroundings as well). A process is called **quasistatic** if it is sufficiently slow such that any intermediate state along the path of the process can be considered to be an equilibrium state. If a process occurs too quickly, the system has no time to internally equilibrate, and so the state variables are no longer well-defined along the path, which automatically renders the process irreversible. The distinction to make here is that while all reversible processes must be quasistatic, not all quasi-static processes are reversible since there is no requirement for a quasi-static process to ensure equilibrium between the system and its surroundings.

If the system is in equilibrium with its surroundings at each point along a process' path, we can construct a continuous function to parametrise the path in terms of state variables, as shown in Fig. 1. Examples of these include isobaric, isochoric, isothermal and adiabatic processes. Specifically for the work, we can consider its dependence on the various state variables $\{X_i\}$ and partition this into the set of **generalised displacements** $\{x_i\}$ and their conjugate **generalised forces** $\{F_i\}$. The work differential is then written as:

$$\mathring{d}W = \sum_i F_i dx_i \quad (1.5)$$

Common examples of force-displacement pairs are the pair of pressure-volume ($-p dV$), magnetic field-magnetisation ($\mathbf{B} \cdot d\mathbf{M}$) and chemical potential-particle number (μdN). Throughout this review, I will typi-

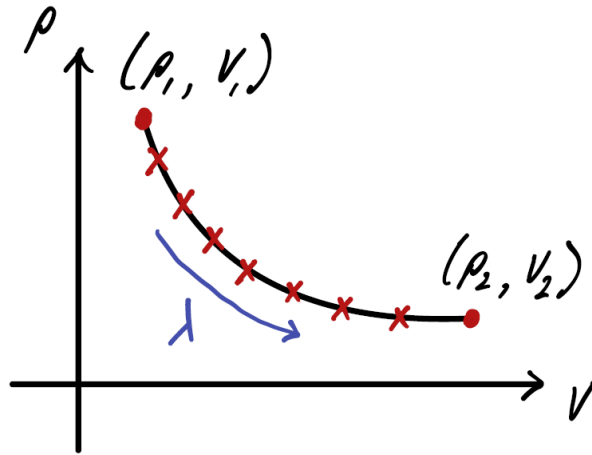


FIGURE 1: Reversible Process. If a process is carried out sufficiently slowly, then its path through state space can be parametrised (through some real-valued parameter λ) in terms of equilibrium values of the state variables at every point along the process.

cally use $dW = -p dV + \mu dN$, with the second term being dropped if it is obvious that the particle number is conserved for the example in question.

1.4 Response Functions

To characterise the macroscopic behaviour of a system, we have to experimentally measure the response of the system to some external perturbation in terms of its thermodynamic state variables. The typical quantity to be measured is known as a thermodynamic **response function**, and it comes in various forms depending on the perturbation being applied and measurement being taken. We first have to emphasise that the ratio of two extensive quantities is intensive, while the ratio of an extensive to an intensive quantity is extensive. The proofs of these are simple, using the basic equation from Def. 0.2 (so they are left for you to try on your own).

With this in mind, we now consider the most familiar response function — the **heat capacity** C_y of a system held at constant y , where y is some thermodynamic state variable. This is formally defined as:

$$C_y = \left(\frac{dQ}{dT} \right)_y \quad (1.6)$$

It should be clear that the heat capacity C_y is an extensive quantity, while the **specific heat** $c_y = C_y/N$ is an intensive quantity. Apart from thermal responses, we can also probe mechanical responses such as the **compressibility** κ , which measures the fractional change in volume of a substance in response to an applied force:

$$\kappa_y = \frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_y \quad (1.7)$$

There are also hybrid thermo-mechanical responses such as the **thermal expansivity** β , which measures the fractional change in volume in response to a change in temperature:

$$\beta_y = \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_y \quad (1.8)$$

In a typical experimental situation, measuring several response functions over a range of parameters allows the experimenter to reconstruct equations of state, which fully characterises the system under study.

1.5 Heat Capacities

Despite our earlier definition of the heat capacity involving a state variable being kept constant, we can (in theory) define the heat capacity for any path as long as the path is differentiable and parametrisable in the relevant state space. We now use this idea to derive the two most commonly used heat capacities C_V and C_p .

Heat Capacity at Constant Volume From the first law, we have:

$$\delta Q = dE + p dV$$

Then, we easily obtain C_V by taking the partial derivative with respect to T , while setting $dV = 0$:

$$C_V = \left(\frac{\delta Q}{\partial T} \right)_V = \left(\frac{\partial E}{\partial T} \right)_V$$

Heat Capacity at Constant Pressure From the first law, we obtain C_p by taking the partial derivative with respect to T at constant pressure:

$$C_p = \left(\frac{\delta Q}{\partial T} \right)_p = \left(\frac{\partial E}{\partial T} \right)_p + p \left(\frac{\partial V}{\partial T} \right)_p$$

Difference in Heat Capacities We will later see that the difference $C_p - C_V$ is a measurable quantity, and is thus of interest to an experimentalist. While that requires some work after we introduce the second law, we can at least try to obtain a simpler expression for this, also using this as a lesson in partial derivative gymnastics. Suppose we have a system with an equation of state in the form $f(p, V, T) = 0$ (e.g. ideal gas, van der Waals gas, etc.). The equation of state sets a constraint on the three variables, indicating that one can always be determined by the values of the other two. We can use this to eliminate the $(\partial E/\partial T)_p$ term in C_p . We first notice that this derivative is canonically obtained from expressing the internal energy E in terms of the pressure and temperature as $E(p, T)$, which gives the total differential:

$$dE = \left(\frac{\partial E}{\partial p} \right)_T dp + \left(\frac{\partial E}{\partial T} \right)_p dT \quad (1.9)$$

It should be noted at this point that it is not possible to write E as a function of $E(p, V, T)$ since the equation of state guarantees that one of these variables is not independent. We are, however, free to choose which variable is dependent at this stage. With this in mind, we now consider the total differential of E when it is written instead as $E(V, T)$:

$$dE = \left(\frac{\partial E}{\partial V} \right)_T dV + \left(\frac{\partial E}{\partial T} \right)_V dT$$

If we are able to convert the dV term here into something that instead contains dp (which is guaranteed to be possible since there is an equation of state linking them), we will obtain an expression for $\partial E/\partial T_p$ in terms of $\partial E/\partial T_V$, successfully eliminating the term we wanted to get rid of. For this, we simply use the total differential of $V(p, T)$ (which exists because of the equation of state):

$$dV = \left(\frac{\partial V}{\partial p} \right)_T dp + \left(\frac{\partial V}{\partial T} \right)_p dT$$

Plugging this into the above equation, we proceed as:

$$\begin{aligned}
dE &= \left(\frac{\partial E}{\partial V}\right)_T \left[\left(\frac{\partial V}{\partial p}\right)_T dp + \left(\frac{\partial V}{\partial T}\right)_p dT \right] + \left(\frac{\partial E}{\partial T}\right)_V dT \\
&= \left(\frac{\partial E}{\partial p}\right)_T dp + \left[\left(\frac{\partial E}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_p + \left(\frac{\partial E}{\partial T}\right)_V \right] dT
\end{aligned}$$

where we use the chain rule in the first term to ‘cancel’ the two appearances⁵ of ∂V . Now, comparing this to the total differential in Eq. 1.9, we easily identify:

$$\left(\frac{\partial E}{\partial T}\right)_p = \left(\frac{\partial E}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_p + \left(\frac{\partial E}{\partial T}\right)_V \quad (1.10)$$

With this result, we can finally insert it into our previous expression for C_p and obtain the heat capacity difference as:

$$C_p - C_V = \left[p + \left(\frac{\partial E}{\partial V}\right)_T \right] \left(\frac{\partial V}{\partial T}\right)_p \quad (1.11)$$

As an example, we can compute this for the ideal gas as:

$$C_p - C_V = \left(\frac{Nk_B T}{V}\right) \frac{V}{T} = Nk_B \quad (1.12)$$

where we note that the derivative $(\partial E/\partial V)_T$ vanishes since the internal energy of the ideal gas is strictly a function of its temperature and not the volume. This result is sometimes also known as **Mayer’s relation**.

1.6 Adiabatic Processes

Definition 1.3: Adiabatic Process

An **adiabatic process** is defined as one that occurs without any exchange of heat between the system and its surroundings. Mathematically, this is written as:

$$\delta Q = 0 \implies \Delta E = \delta W \quad (1.13)$$

which implies that any change in the internal energy is due to the work alone. This typically occurs when a process is so rapid that it is completed on a much shorter time scale than is required for thermal equilibration to occur between the system and its surroundings.

We can evaluate the defining relation for an adiabatic process⁶ in the case of an ideal gas, producing curves known as **adiabats**. Since the ideal gas equation of state is of the form $f(p, V, T) = 0$, we expect that the imposition of the adiabatic constraint results in a single-parameter trajectory in this state space (i.e. adiabats are just a family of curves). We start from the fundamental relation as usual:

$$dE = \delta Q - p dV \implies dE + p dV = 0$$

⁵Mathematicians would burn me at the stake for saying this, but you know exactly what I mean when I say this.

⁶You may notice that this is in complete contradiction with the idea of an adiabatic process in the context of classical and quantum mechanics (look up adiabatic invariants or geometric phase), where an adiabatic classical/quantum process which modifies the Hamiltonian takes place slowly enough such that the system beginning in the ground state always remains in the ground eigenstate. This is due to the unfortunate coincidence that old quantum theory used the term adiabatic to describe this process, even though it has nothing to do with the thermodynamic definition of adiabatic.

where we have set $\dot{d}Q = 0$ for the adiabat. Now, we know that the internal energy of an ideal gas is a function of its temperature alone, and so we can write:

$$E = E(T) \implies C_V = \left(\frac{\partial E}{\partial T} \right)_V = \frac{dE}{dT}$$

which can be used to substitute the differential of the internal energy. We then proceed to try and integrate for the adiabat:

$$\begin{aligned} 0 &= C_V dT + \frac{Nk_B T}{V} dV \\ \frac{dT}{T} &= -\frac{Nk_B}{C_V} \frac{dV}{V} \\ T &= V^{-Nk_B/C_V} + c, \quad c \in \mathbb{R} \\ c &= TV^{C_p/C_V - 1} \\ \implies c &= PV^\gamma \end{aligned} \tag{1.14}$$

In the third line, the constant c is simply a real constant of integration which differs on all subsequent lines by factors. In the fourth line, we use Mayer's relation from Eq. 1.12 to introduce the constant pressure heat capacity. Finally, we define $\gamma = C_p/C_V$ as the **heat capacity ratio** in the last line, which gives us the defining equation for an adiabat.

2 The Second Law of Thermodynamics

From the first law, we have a picture of thermodynamics that is consistent with the conservation of energy. We have not, however, refined this picture into one that is independent of a system's history. Specifically, our working definition of heat involves the path variable Q and so it is difficult to describe the equilibrium state functions of a system at any point in time as it navigates the state space. For this, we will need to motivate and develop a state variable descriptor of heat, which will ultimately lead us to the second law of thermodynamics. It should be noted here that thermodynamics does not actually provide a derivation for the second law, and it is merely a self-consistent framework that fits into the rest of thermodynamics. The only way to *derive* the second law is to use statistical mechanics, but I will not do so here since there are many subtleties involved that present countless points of confusion and argument. Instead, we will study the dynamics of heat and motivate the construction of a state variable from there.

2.1 Heat Engines and Refrigerators

Early studies of thermodynamics involved developing a description for the flow of heat, and eventually led to principles governing the driven flow of heat in devices that have now become indispensable for everyday life, such as heat engines and refrigerators (or heat pumps). At an abstract level, a **heat engine** is a cyclic process \mathcal{E} in a system that absorbs heat $|Q_{\text{in}}|$, rejects heat $|Q_{\text{out}}|$, and performs a positive amount of work $|W|$. Here, we define Q and W to be positive for an inflow of energy into the system. Now, we apply the first law to this engine to obtain:

$$Q = |Q_{\text{in}}| - |Q_{\text{out}}| = -W$$

We can also define the efficiency η of the heat engine as:

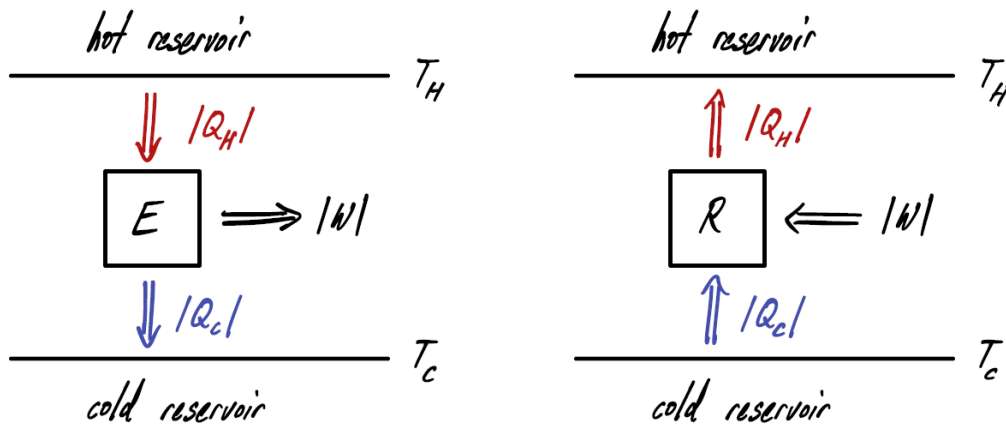


FIGURE 2: Heat Engines and Refrigerators. (Left) A heat engine \mathcal{E} is any cyclic process in a system that accepts heat $|Q_H|$ from a hot reservoir at temperature T_H and outputs some useful work $|W|$, rejecting waste heat $|Q_C|$ to a cold reservoir at T_C in the process. (Right) A refrigerator \mathcal{R} (or heat pump) is any cyclic process in a system that removes heat $|Q_C|$ from a cold reservoir at temperature T_C and deposits heat $|Q_H|$ into a hot reservoir at T_H , requiring some input work $|W|$ in the process.

$$\eta = \frac{\text{work out}}{\text{heat in}} = 1 - \frac{|Q_{\text{out}}|}{|Q_{\text{in}}|} \quad (2.1)$$

which also constrains the efficiency to lie in the range $\eta \in [0, 1]$ by the conservation of energy. In general, two-temperature devices such as the heat engine are cyclic devices working between two heat reservoirs. In the context of thermodynamics, we consider a reservoir to simply be an external system that contains such a large amount of matter and energy that the addition or removal of heat/work by our cyclic process does not change its temperature. Heat engines thus specifically accept heat $|Q_H|$ from a hot reservoir at temperature T_H , perform some useful work $|W|$ and reject waste heat $|Q_C|$ at temperature T_C , with the imposition that $T_H > T_C$. With these formalities in place, we can also define the heat pump, or **refrigerator** to be a cyclic process which accepts heat $|Q_C|$ from a cold reservoir at T_C and accepts work $|W|$, then deposits heat $|Q_H|$ into a hot reservoir at temperature T_H . These devices are shown pictorially in Fig. 2. The efficiency of a refrigerator is then defined as⁷:

$$\eta = \frac{\text{heat in}}{\text{work in}} = 1 - \frac{|Q_C|}{|W|} \quad (2.2)$$

2.2 The Second Law

Please indulge me for a moment, for I would be remiss if I did not share perhaps my favourite description of the second law, written by Elliott H. Lieb and Jakob Yngvason [6].

“The second law of thermodynamics is, without a doubt, one of the most perfect laws in physics. Any *reproducible* violation of it, however small, would bring the discoverer great riches as well as a trip to Stockholm. The world’s energy problems would be solved at one stroke. It is not possible to find any other law (except, perhaps, for super selection rules such as charge conservation) for

⁷The efficiency of a heat pump is defined slightly differently to that of a refrigerator despite performing the same physical job. This is because the heat pump’s purpose is to supply the hot reservoir with heat, while the refrigerator’s job is to remove heat from the cold reservoir. The efficiency of a heat pump is defined as: $\eta = \text{heat out}/\text{work in} = |Q_H|/|W|$.

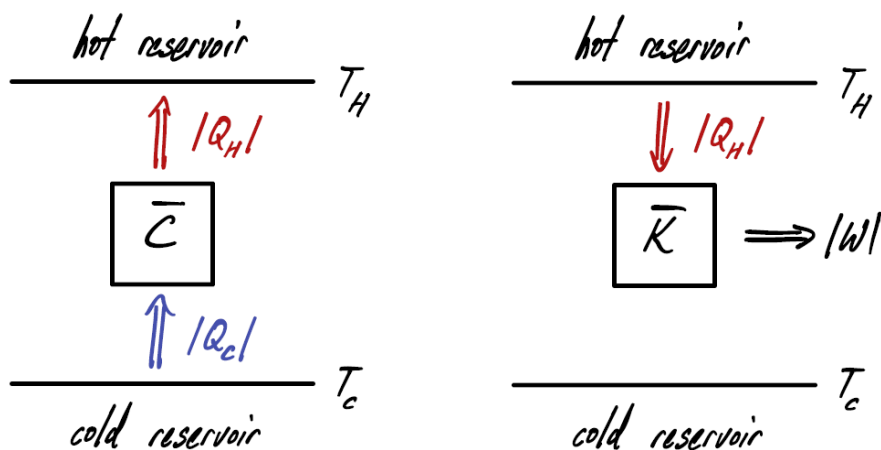


FIGURE 3: Impossible Devices from the Clausius and Kelvin–Planck Statements. (Left) A *Clausius violator* is an impossible device which violates the Clausius statement of the second law, by removing heat from a cold reservoir and depositing it entirely as heat in a hot reservoir without any input work. (Right) A *Kelvin violator* is an impossible device which violates the Kelvin–Planck statement of the second law, by accepting heat $|Q_H|$ and converting it entirely to work with unit efficiency, generating no waste heat in the process.

which a proposed violation would bring more scepticism than this one. Not even Maxwell’s laws of electricity or Newton’s law of gravitation are so sacrosanct, for each has measurable corrections coming from quantum effects or general relativity. The law has caught the attention of poets and philosophers and has been called the greatest scientific achievement of the nineteenth century. Engels disliked it, for it supported opposition to dialectical materialism, while Pope Pius XII regarded it as proving the existence of a higher being (Bazarow, 1964, Section 20).” — Lieb & Yngvason (1999) [6].

Historically, the first studies of heat transfer and efficiency of cyclic processes were conducted by Sadi Carnot and published in his book — *Reflections on the Motive Power of Fire* [3]. Tragically, like many other young adults in France during the 1800s, he succumbed to cholera at the young age of 36, and many of his writings were buried along with him due to fear of the contagious nature of the disease. As a result, it is believed that progress in fundamental thermodynamics was delayed by at least a decade before Rudolf Clausius finally formulated his statement of the second law in 1850. Towards the end of the 19th century, Lord Kelvin⁸ provided a similar formulation together with Max Planck, now known as the Kelvin–Planck statement of the second law. Immediately after this work, Planck turned his attention to the problem of blackbody radiation, and developed his now renowned Planck distribution of blackbody radiation which solved the ultraviolet catastrophe. For this reason, Carnot, Clausius, Kelvin and Planck are frequently dubbed as the early fathers of thermodynamics⁹. In this section, we will study the Clausius and Kelvin–Planck statements, and show that they are equivalent.

⁸His full name was William Thomson, 1st Baron Kelvin.

⁹Of course, this list of names is not complete. If I were to elaborate on the complete history of thermodynamics, this review would be a full-length novel (and Netflix series) that I am not currently being paid enough to write.

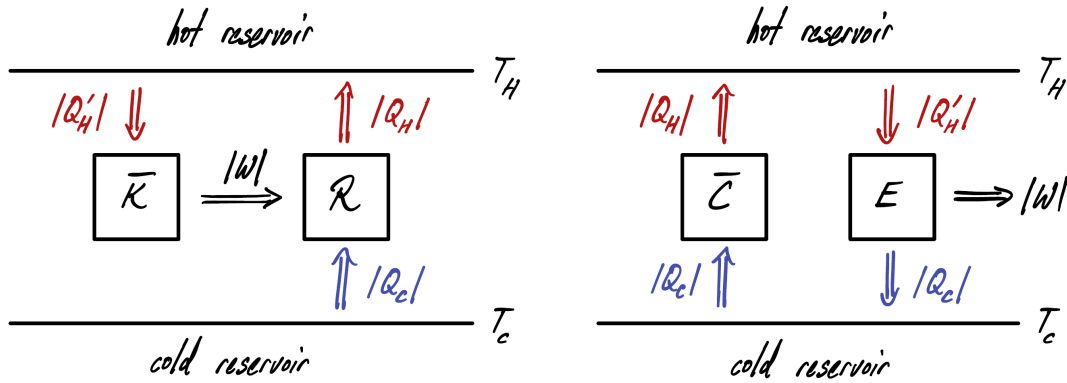


FIGURE 4: Constructions of Clausius and Kelvin Violators for Equivalence. (Left) The composite device $(\bar{K} \circ \mathcal{R}) \equiv \mathcal{F}$ formed by combining a *Kelvin violator* and a refrigerator can be shown to be equivalent to a *Clausius violator*, by spontaneously driving heat from a cold reservoir to a hot reservoir with no work input. (Right) The composite device $(\bar{C} \circ \mathcal{E}) \equiv \mathcal{G}$ formed by combining a *Clausius violator* with a heat engine can be shown to be equivalent to a *Kelvin violator*, by converting heat from a hot reservoir into useful work with perfect efficiency, generating no waste heat.

Definition 2.1: Clausius and Kelvin–Planck Statements

The Clausius Statement P_C : There is no process \bar{C} whose only effect is to accept heat from a cold reservoir at T_C and transfer it to a hot reservoir at T_H , with no work input. Equivalently, heat cannot be (spontaneously) transferred from a cold reservoir to a hot reservoir without work.

The Kelvin Statement P_K : There is no process \bar{K} whose only effect is to accept heat from a single heat reservoir and convert it entirely into work. Equivalently, a heat engine must always reject some waste heat.

These impossible processes/devices are shown pictorially in Fig. 3.

To prove the equivalence of these two statements¹⁰ ($P_C \Leftrightarrow P_K$), we need to show both $P_C \Rightarrow P_K$ and $P_K \Rightarrow P_C$. For this, we will show each premise using proofs by contraposition.

Theorem 2.1. *The Clausius statement P_C and Kelvin–Planck statement P_K are equivalent.*

Proof. We will first show the contrapositive $\neg P_K \Rightarrow \neg P_C$. Suppose we have a *Kelvin violator* \bar{K} which accepts $|Q'_H|$ from a hot reservoir at T_H and converts it entirely into work $|W|$. We place this alongside a refrigerator \mathcal{R} operating between a cold reservoir at T_C and the same hot reservoir, such that it accepts heat $|Q_C|$ from the cold reservoir, takes in work $|W|$ from \bar{K} and deposits heat $|Q_H|$ into the hot reservoir, as shown in Fig. 4. The composite device $(\bar{K} \circ \mathcal{R}) \equiv \mathcal{F}$ now has the net effect of removing heat $|Q_C|$ from the cold reservoir and depositing heat $|Q_H| - |Q'_H|$ at the hot reservoir, with no work input! Thus, the composite system \mathcal{F} is equivalent to \bar{C} , a *Clausius violator*!

Next, we show the contrapositive $\neg P_C \Rightarrow \neg P_K$. Suppose we have a *Clausius violator* \bar{C} which removes heat $|Q_C|$ from a cold reservoir and deposits it entirely as heat $|Q_H| = |Q_C|$ into the hot reservoir. Then, we consider a heat engine \mathcal{E} that accepts heat $|Q'_H|$ from the hot reservoir and rejects heat $|Q_C|$ at the cold reservoir,

¹⁰It is hard to prove the absolute truth of these statements in a purely thermodynamic fashion, as the typically used proofs in undergraduate courses now simply resort to statistical mechanics arguments. When these statements were introduced, it was simply widely accepted that the Clausius statement was true by intuition that heat does not spontaneously leave a colder body and flow into a hotter body. During the era of these developments, showing the truth of a statement of the second law generally involved one showing that their conjecture was somehow equivalent to the Clausius statement.

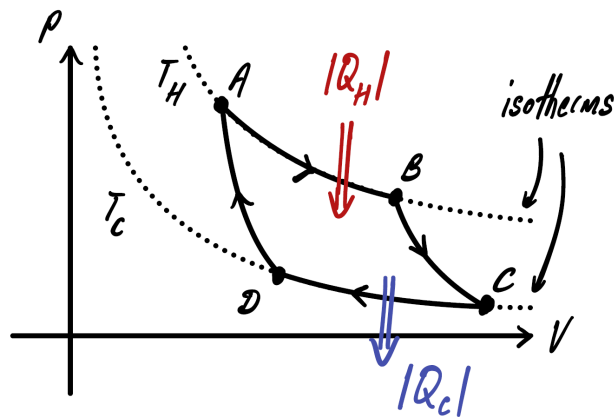


FIGURE 5: The Carnot Cycle. A Carnot engine undergoing a Carnot cycle involves a sequence of reversible processes which maximise the efficiency of an engine operating between two reservoirs. Heat is transferred into the engine through isothermal expansion at the hot reservoir, before the system is adiabatically cooled to the cold reservoir. Heat is then rejected at the cold reservoir through isothermal compression, before it is finally adiabatically heated back to the hot reservoir. The total work done in this process is easily obtained by integrating to find the area of the region enclosed by the process in the p - V state space.

performing work $|W|$ in the process, as shown in Fig. 4. The composite device $(\bar{C} \circ \mathcal{E}) \equiv \mathcal{G}$ now has the net effect of absorbing heat $|Q'_H| - |Q_H|$ from the hot reservoir and performing work $|W|$, with no waste heat generated! Thus, the composite system \mathcal{G} is equivalent to \bar{K} , a *Kelvin violator*!

Combining the two contrapositive statements, we conclude that the Clausius and Kelvin–Planck statements are indeed equivalent. \square

2.3 Carnot’s Theorem

Now that we have working statements of the second law in place, we will take a step back to visit Carnot’s initial work that was fundamental in motivating Clausius’ and Kelvin’s statements and later formulations of the second law. To understand Carnot’s brilliance, we must first dissect his definition of a perfect heat engine. A **Carnot engine** is an idealised system which is cyclic and reversible, with all heat exchanges taking place only at either a hot reservoir T_H or a cold reservoir T_C .

Definition 2.2: Carnot Cycle

For two temperatures $T_C < T_H$, the **Carnot cycle** is defined as a sequence of four processes (shown in Fig. 5), starting from the hot reservoir at T_H :

1. $A \rightarrow B$: Isothermal expansion at T_H
2. $B \rightarrow C$: Adiabatic cooling from T_H to T_C
3. $C \rightarrow D$: Isothermal compression at T_C
4. $D \rightarrow A$: Adiabatic heating from T_C to T_H

We note that the heat transfer along adiabatic paths is zero $\Delta Q = 0$, and thus heat is only exchanged in

this process along the isotherms at the two reservoirs. The efficiency of a Carnot engine, also known as the **Carnot efficiency**, is thus obtained as:

$$\eta_{CE} = 1 - \frac{|Q_C|}{|Q_H|} = 1 - \frac{T_C}{T_H} \quad (2.3)$$

The second equality here is a more general result from Carnot, which we will shortly explore. For now, we will demonstrate the truth of the second equality for a Carnot engine operating using an ideal gas as its medium. We first use the fact that the internal energy of an ideal gas solely depends on its temperature, so the internal energy is constant during an isothermal process. Thus, we can integrate the equation of state to determine the heat transfer taking place during an isothermal process from initial volume V_i to final volume V_f :

$$\begin{aligned} 0 &= \text{d}Q + \text{d}W \\ \text{d}Q &= -Nk_B T \frac{\text{d}V}{V} \\ \implies Q_{\text{in}} &= Nk_B T \log\left(\frac{V_f}{V_i}\right) \end{aligned}$$

where the sign is flipped because the right hand side of the last equation details the work done *by* the system (as opposed to the work done *on* the system). Using this, we now identify the input and output heats of a Carnot engine as:

$$|Q_{\text{in}}| = Q_{A \rightarrow B} = Nk_B T_H \log\left(\frac{V_B}{V_A}\right), \quad |Q_{\text{out}}| = -Q_{C \rightarrow D} = -Nk_B T_C \log\left(\frac{V_D}{V_C}\right)$$

The efficiency is now written as:

$$\eta_{CE} = 1 - \frac{|Q_{\text{out}}|}{|Q_{\text{in}}|} = 1 - \frac{T_C}{T_H} \log\left(\frac{V_B V_D}{V_A V_C}\right)$$

To eliminate the appearances of the volumes, we return to the adiabat equation from Eq. 1.14 to write:

$$PV^\gamma = \text{const.} \implies TV^{\gamma-1} = \text{const.}$$

Plugging in the values for the two adiabat curves in the Carnot cycle, we arrive at:

$$\begin{cases} T_B V_B^{\gamma-1} = T_C V_C^{\gamma-1} \\ T_D V_D^{\gamma-1} = T_A V_A^{\gamma-1} \end{cases} \implies \frac{V_C}{V_B} = \frac{V_D}{V_A}$$

Finally, substituting this into the efficiency produces the second equality as required, demonstrating that the efficiency of the Carnot engine operating using an ideal gas depends only on the temperature ratio between the two reservoirs. Carnot was able to prove this relation for all working media, but we will not do so here as it is rather involved. We now proceed to state and prove **Carnot's theorem**, which reads as follows:

Theorem 2.2. *No engine operating between two reservoirs can be more efficient than a Carnot engine operating between the same two reservoirs, regardless of the working medium¹¹.*

¹¹It should be noted here that the theorem is not necessarily satisfied if the engine operates between more than two reservoirs instead. That is, if heat exchange occurs in multiple stages between different reservoirs, one can exceed the Carnot efficiency dictated by the temperatures of the two reservoirs at the temperature extrema.

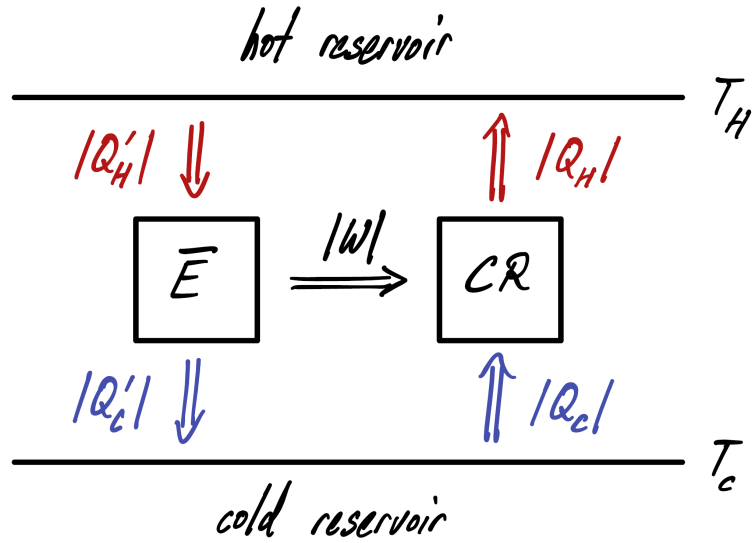


FIGURE 6: Proof of Carnot's Theorem by Contradiction. To prove Carnot's theorem, we consider a Carnot engine running in reverse to effectively act as a Carnot refrigerator CR . The required work for its operation is then provided by a hypothetical super-efficiency engine \bar{E} which accepts less heat from the hot reservoir than CR deposits. Applying the first law then easily shows that the composite device formed by $(\bar{E} \circ CR)$ is equivalent to a *Clausius violator*, forcing a contradiction.

Proof. Suppose we have a Carnot engine operating between two reservoirs at $T_C < T_H$, accepting heat $|Q_H|$ from the hot reservoir and rejecting heat $|Q_C|$ at the cold reservoir, providing $|W|$ useful work in the process. We now reverse the engine to obtain a Carnot refrigerator with the same heats at the two reservoirs, keeping the definition of the efficiency as that of a heat engine for now. That is, we define the efficiency as $\eta_{CR} = \text{workin}/\text{heatout}$. Now, for sake of contradiction, assume we have an impossible super-efficient engine \bar{E} which operates between the same two reservoirs, which accepts heat $|Q'_H|$ from the hot reservoir and rejects heat $|Q'_C|$ at the cold reservoir. This also provides the required work $|W|$ for the Carnot refrigerator in the process, but with an efficiency $\eta_{SE} > \eta_{CR}$. It should be clear that since \bar{E} is able to provide the required work $|W|$ at higher efficiency than CR , the input heat $|Q'_H|$ to \bar{E} must be less than the deposited heat $|Q_H|$ by CR (this is clear if one looks at the definition of the efficiency). These devices are shown in Fig. 6. Now, since each device is cyclic in operation, we can apply the first law to each of them individually knowing that $\Delta E = 0$ to write:

$$|W| = |Q'_H| - |Q'_C| = |Q_H| - |Q_C| \implies |Q_H| - |Q'_H| = |Q_C| - |Q'_C| > 0$$

This tells us that, at the hot reservoir there is a net inflow of energy, while at the cold reservoir there is a net outflow of energy. With this, it becomes clear that the composite device $(\bar{E} \circ CR)$ is equivalent to a *Clausius violator*! Thus, by contradiction, we conclude that such a super-efficient engine \bar{E} cannot exist, and thus the maximal efficiency for any engine operating between two reservoirs is given by the Carnot efficiency. \square

Using the above conclusion, it is also trivial to prove that *all* Carnot engines operating between the same two reservoirs must have exactly the same efficiency¹².

¹²This is a simple proof by relabelling.

2.4 Clausius' Inequality

Carnot's observations of the theoretical maximum efficiency of heat engines went largely unlinked to the second law of thermodynamics for some time after his death, until Rudolf Clausius observed that one could develop a path-independent quantity through Carnot's theoretical construction. The actual formulation of the concept of entropy began with Clausius' studies of bounds that one could place on the nature of thermodynamic processes altogether. This ultimately manifested in what is now known as **Clausius' inequality**, which is a remarkably powerful statement on the extent of heat exchange allowed to occur within cyclic processes. The derivation of this inequality is rather involved, so I will try to provide as much elaboration and visualisation as possible.

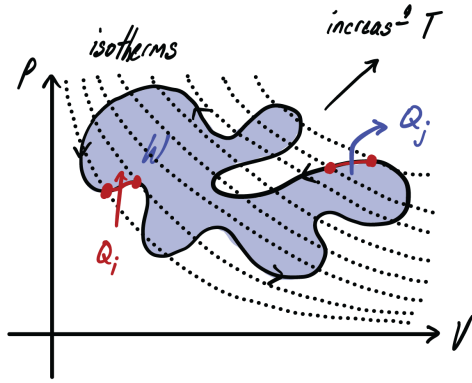


FIGURE 7: Arbitrary Process in State Space as a Sequence of Infinitesimal Heat Exchanges. Since we are considering a cyclic process in state space, this must be a curve parametrised by a single parameter. Thus, we can consider the path as a function of this parameter broken into infinitesimal pieces, moving from one value T_k to the next T_{k+1} . In the p - V state space, this is visualised as the curve intersecting various isotherms over the course of the cyclic process. If the curve moves from a lower isotherm to a higher one, corresponding to increasing temperature, we deduce that heat must have entered the system at that step (and vice versa). In the diagram, this corresponds to Q_i being a heat input, while Q_j is a heat output. The total area enclosed by the cyclic process is the work done on the system.

combining the engine and all Carnot cycles. The heat exchanges $\{Q_i\}$ are now internal to \mathcal{F} , and so we can ignore them altogether¹⁴. Thus, the net effect of \mathcal{F} is to exchange heat $\sum_i Q'_i$ with a single heat reservoir at T_0 and

Suppose we have an engine \mathcal{E} which undergoes some reversible and cyclic process as shown in Fig. 7. We can define a temperature (state variable) at each point in the cycle, then consider the heat exchange necessary for \mathcal{E} to successfully transit from each point T_k to the next point T_{k+1} along its path¹³. We can represent each of these heat exchanges as \mathcal{E} receiving heat from or donating heat to a Carnot cycle \mathcal{C}_i at each temperature T_i , as it performs the i -th temperature transit. Each Carnot cycle may either be an engine or refrigerator, depending on the direction of the temperature change required by the cycle undertaken by \mathcal{E} . Now, each Carnot cycle thus exchanges heat $|Q'_i|$ with some reservoir T_0 , performs or receives work $|W_i|$ then exchanges heat $|Q_i|$ with \mathcal{E} at T_i , bringing \mathcal{E} to T_{i+1} . As \mathcal{E} exchanges heat with each Carnot cycle and undergoes a full cycle, it either receives or performs total work $|W_0|$.

With our setup in place, we now drop all absolute-value functions and make a careful attempt to work things out with a consistent sign convention. Specifically, we define Q'_i and Q_i as positive if they enter \mathcal{C}_i , W_i as positive if it enters \mathcal{C}_i , and W_0 as positive if it enters \mathcal{E} . We then apply the first law to \mathcal{E} and all Carnot cycles $\{\mathcal{C}_i\}$:

$$\Delta E_{\mathcal{E}} = - \sum_i Q_i + W_0 = 0 \quad (\text{applied to } \mathcal{E})$$

$$\Delta E_i = Q_i + Q'_i + W_i = 0 \quad (\text{applied to } \mathcal{C}_i)$$

where the right hand sides trivially vanish since all processes listed are reversible and cyclic, and thus the internal energy must not change over a single cycle. Now, we consider the composite device $(\mathcal{E} \circ \bigcup_i \mathcal{C}_i) \equiv \mathcal{F}$ formed by combining

¹³Knowing that the internal energy of the system will depend on its temperature, the first law tells us that either heat or work must be provided for the temperature to change.

¹⁴Another way of seeing this is that for every Q_i entering \mathcal{C}_i , we have a corresponding $-Q_i$ entering \mathcal{E} .

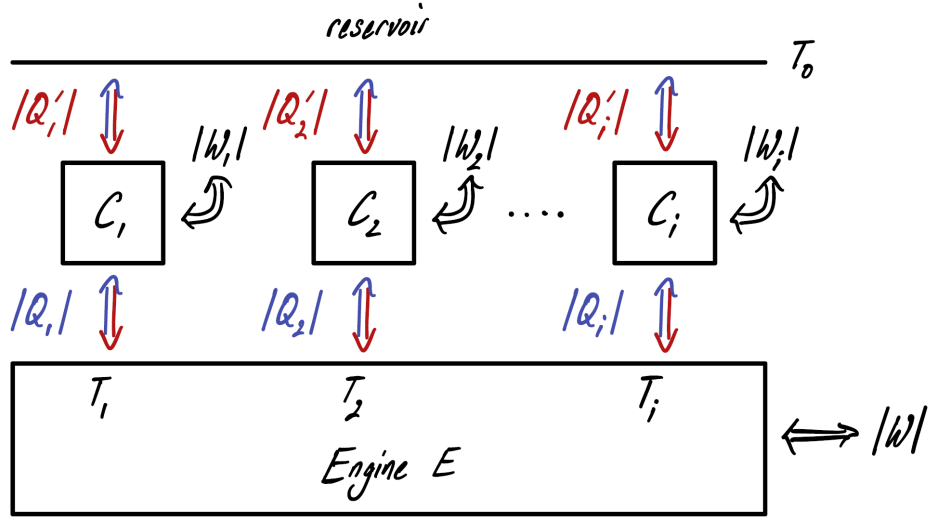


FIGURE 8: Proof of Clausius' Inequality. In proving this inequality, we consider a cyclic process undertaken by an engine \mathcal{E} , with each step enabled by heat exchange $|Q_i|$ with some Carnot engine/refrigerator at temperature T_i . Each Carnot cycle then exchanges heat $|Q'_i|$ with common reservoir at T_0 , performing or receiving work $|W_i|$ in the process. The engine \mathcal{E} as a whole then has to either have zero work output or solely work input $W_0 \geq 0$, otherwise the composite device formed by $(\mathcal{E} \circ \bigcup_i C_i) \equiv \mathcal{F}$ would constitute a *Kelvin violator*. Each heat arrow is also bi-coloured red and blue here to denote the fact that heat exchange can occur in either direction, depending on the change of T_i required at the i -th step of \mathcal{E} 's cyclic process.

perform total work $(W_0 + \sum_i W_i)$. At this point, we should note that if heat enters \mathcal{F} and work leaves (there is no other reservoir for heat to be rejected at), then we have a *Kelvin violator*! Knowing that we're headed in this direction, we apply the first law to \mathcal{F} :

$$\Delta E_{\mathcal{F}} = \sum_i Q'_i + W_0 + \sum_i W_i = 0$$

By the condition established above using the Kelvin-Planck statement, we know that $\sum_i Q'_i > 0$ is forbidden. We thus have the restriction that $\sum_i Q'_i \leq 0$. We now constrain the heats exchanged by each of the individual Carnot cycles using the Carnot efficiency:

$$-\frac{Q_i}{Q'_i} = \frac{T_i}{T_0} \implies Q'_i = \frac{T_0}{T_i}(-Q_i) \implies \sum_i \frac{T_0}{T_i}(-Q_i) \leq 0$$

where the initial negative sign is due to the fact that Q'_i enters C_i while Q_i leaves it. Factoring out the common reservoir temperature T_0 , we obtain Clausius' inequality as [2]:

$$\sum_i \frac{(-Q_{i,C_i})}{T_i} = \sum_i \frac{Q_{i,\text{sys}}}{T_i} \leq 0 \quad (2.4)$$

where we note that $-Q_i$ entering C_i is equivalent to Q_i entering \mathcal{E} , our system of interest. This tells us that the heats $\{Q_i\}$ exchanged by a system at some temperatures T_i must sum according to the above inequality, for any cyclic process. Specifically in the case of a reversible and cyclic process, we would have:

$$\begin{cases} \sum_i \frac{Q_i^{\text{forward}}}{T_i} \leq 0 \\ \sum_i \frac{Q_i^{\text{backward}}}{T_i} \leq 0 \end{cases} \implies \sum_i \frac{Q_i}{T_i} = 0$$

The equality is implied since the heats from the backward process are flipped in sign with respect to the heats from the forward process, so the combined inequality implies that they must sum to exactly zero. The distinction between reversible and irreversible processes here was subtle, and only came up right at the end. What we did was to start with the assumption of reversibility, but this is not really necessary if we were to just allow for the cyclic process to be quasistatic so an internal temperature can be defined for each of the heat exchanges between \mathcal{E} and \mathcal{C}_i . We thus see that the inequality strictly holds for an irreversible process, with equality only occurring in the special case of reversibility, where the system is in equilibrium with the surroundings at every step of the process.

2.5 Entropy as a State Function

Having proven Clausius' inequality, we now see that the special case of equality for reversible processes looks a lot like the fundamental theorem of calculus for line integrals, otherwise known as the **gradient theorem**. This is simply begging to be turned into a state function, and we will start doing so by dividing the entire reversible cyclic process into N steps with paths C_i , where the entire cyclic path is traced out by the union of all small steps as $C \equiv \bigcup_i C_i$. Now, each heat exchange Q_i can be written in terms of a line integral over the path variable:

$$Q_i = \int_{C_i} \mathrm{d}Q$$

With this, we have the entire sum of paths as:

$$\begin{aligned} \sum_{i=1}^N \frac{Q_i}{T_i} &= \sum_{i=1}^N \frac{1}{T_i} \int_{C_i} \mathrm{d}Q \\ &\approx \sum_{i=1}^N \int_{C_i} \frac{\mathrm{d}Q}{T_i} \end{aligned}$$

where we pull the temperature into the line integral under the assumption that each path is sufficiently short (N is sufficiently large). Continuing this limit, we thus see:

$$\lim_{N \rightarrow \infty} \sum_{i=1}^N \frac{Q_i}{T_i} = \oint \frac{\mathrm{d}Q}{T} = 0 \quad (2.5)$$

This is an incredible result, as it finally tells us that the combined differential $\mathrm{d}Q/T$ is an exact differential, and thus independent of the path traversed in state space! We give the variable formed by this exact differential the symbol S , and this is what we have come to love (and hate) as **entropy**:

$$\mathrm{d}S = \frac{\mathrm{d}Q}{T} \implies \mathrm{d}Q = T \mathrm{d}S \quad (2.6)$$

Finite changes in entropy are then quantified by using the path integral as:

$$\Delta S = S_f - S_i = \int_C \frac{\mathrm{d}Q}{T}$$

2.6 The (Practical) Second Law

We finally have all the tools we need to derive the second law of thermodynamics in the form that we have all been used to hearing. For this, we first consider two states of a system A and B such that there are two processes linking them:

1. $A \rightarrow B$: Irreversible
2. $B \rightarrow A$: Reversible

with the entire cycle $A \rightarrow B \rightarrow A$ being cyclic. Then, by Clausius' inequality, we have (for the entire cycle):

$$\sum_i \frac{Q_i}{T_i} = \sum_j \frac{Q_j^{\text{irrev}}}{T_j} + \sum_k \frac{Q_k^{\text{rev}}}{T_k} \leq 0$$

Notice that for the reversible process, each point has a well-defined temperature since each point is an equilibrium state and the curve is continuous in the state space. With the continuity criterion satisfied, we can proceed to infinitesimally subdivide the reversible process and invoke the differential limit:

$$\sum_k \frac{Q_k^{\text{rev}}}{T_k} \Rightarrow \int_{C_{B \rightarrow A}} \frac{\mathrm{d}Q}{T} = S_A - S_B$$

Now, we do the somewhat obvious step of reversing the reversible process. This will flip the signs of each Q_k^{rev} , so the inequality becomes:

$$\sum_j \frac{Q_j^{\text{irrev}}}{T_j} \leq \int_{C_{A \rightarrow B}} \frac{\mathrm{d}Q}{T} = S_B - S_A = \Delta S$$

Finally, if we contain the entire system in question within rigid, insulating walls, this constitutes a perfect isolated system. Then, there cannot be any heat exchange between the system and its surroundings, so we must have $Q_j^{\text{irrev}} = 0$ at each step. This finally gives us the **second law of thermodynamics**¹⁵:

Definition 2.3: Second Law of Thermodynamics

The entropy of an isolated system can never decrease over the course of its evolution. Mathematically, this is written as:

$$\Delta S = \int \frac{\mathrm{d}Q^{\text{rev}}}{T} \geq 0 \quad (2.7)$$

with equality only obtained for a reversible process.

As stated previously, the only real isolated system (that we possibly know of) is the Universe as a whole. This is the origin of the statement that the entropy of the Universe can never decrease. In a typical process that involves heat exchange between a system and a reservoir, the spontaneous flow of heat from the system to the reservoir constitutes an irreversible process, since this is an ongoing process of equilibration.

¹⁵I vividly remember this moment as an undergraduate, when I saw the derivation of the second law completed from start to finish (the last nine or so pages) over the course of two hours in a lecture. It was incredibly beautiful and poignant in some strange manner that I can't really describe, formalising the intuitive ideas that we have of energy and information being irretrievably lost in microscopic degrees of freedom, never to return as work. I hope you felt some of that here. It also doesn't help that I'm currently listening to the final track of Muse's album *The Second Law* as I write this section, the track titled *The Second Law: Isolated System*; the emotive power of physics is sometimes felt in moments like these.

2.7 The Principle of Entropy Maximisation

With an understanding of entropy as applied to idealised constructions in deriving the second law, it is perhaps wise to apply it to some simple systems to develop some intuition of its mathematical structure. We first consider the implications that the second law has for the idea of reversibility in isolated systems. For this, consider an isolated system consisting of a gas in a box. The gas is constrained to reside within some corner of the box (not occupying the full volume of the box), held in place by some internal constraint such as a non-porous and rigid membrane as shown in Fig. 9.

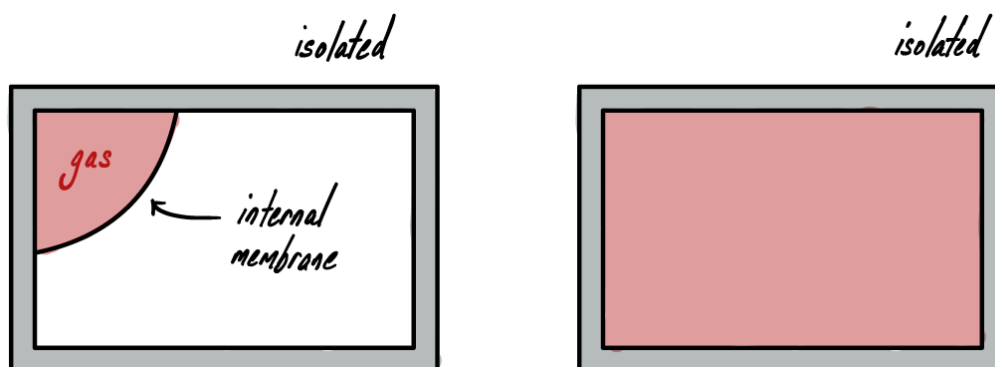


FIGURE 9: Free Expansion of an Isolated Gas. Starting with a gas constrained (using some internal membrane) to reside in a corner of an isolated box (left), we abruptly remove the membrane and allow the gas to freely expand into the vacuum of the remainder of the box. The final state of the gas is an equilibrium state in which the gas occupies the full volume of the box (right), and involves the entropy of the gas increasing due to the free expansion. The reverse of this process, however, cannot occur spontaneously.

In this initial state, the constrained gas is at equilibrium with some parameters (E_i, V_i, N_i) . Now we suppose that the membrane is abruptly removed, allowing the gas to freely expand to occupy the full volume of the box, in a process known as **free expansion** (or sometimes also Joule expansion, not to be confused with the Joule-Thomson process). The gas will now attain some new equilibrium with parameters (E_f, V_f, N_f) . Since the system is isolated, we expect that energy is conserved ($E_i = E_f$) and particle number is conserved ($N_i = N_f$). Furthermore, the free expansion of the gas into a vacuum implies that no work is performed by the gas, nor is work done on the gas by any other part of the system. Application of the first law to these premises then leads us to the conclusion that this entire process must have been adiabatic, so $\Delta Q = 0$.

Now, suppose we wanted to study the reverse process — the *free compression* of an isolated gas. In order to bring the fully expanded gas back into a corner, we have to perform some work on it to compress it against its pressure. The first law (with the imposition that $\Delta E = 0$ since this is isolated) then tells us that heat must be transferred in the process. Specifically, work must be done on the gas and an equal amount of heat must be extracted from it to keep the internal energy constant. However, since this is occurring from a single source, this is equivalent to a *Kelvin violator* where heat is extracted from a reservoir and converted entirely to work with unit efficiency¹⁶. This leads us to the conclusion that the reverse process is not possible, and so the free expansion process is an irreversible one. How did this happen?

We previously established that $dS = \delta Q^{\text{rev}}/T \geq 0$ in our statement of the second law, and this process of free expansion is an ideal example of when equality is no longer achieved. Specifically, the fact that the gas is at each point in its expansion in a non-equilibrium state (starting from the instant the membrane vanishes) allows

¹⁶There is a subtle point to note here — the fact that the internal energy remains constant implies that it must not change in temperature. Thus, the isothermal compression process is akin to extracting heat from a reservoir at fixed temperature.

for the total entropy change in the expansion process to be positive, without any heat transfer occurring. This can be exactly computed for the ideal gas as:

$$\Delta S = \int dS = \int_{V_i}^{V_f} \frac{p dV}{T} = \int_{V_i}^{V_f} \frac{Nk_B dV}{V} = Nk_B \log \left(\frac{V_f}{V_i} \right) > 0 \quad (2.8)$$

Once in the final equilibrium state, the parameters of the gas are once again well-defined and unchanging, leading to $dS = 0$ at that point. Seeing that the isolated gas increases its entropy in expansion until it reaches an equilibrium at which point it is unchanged, we conclude that entropy is maximised at equilibrium in an unconstrained isolated system. This is the principle of **entropy maximisation**. This tells us that, in the equilibrium state, the partial derivatives of the entropy with respect to internal parameters satisfy:

$$\left(\frac{\partial S}{\partial X} \right)_E = 0, \quad \left(\frac{\partial^2 S}{\partial X^2} \right)_E < 0 \quad (2.9)$$

where X is any extensive parameter apart from the internal energy (typically in the form of some mechanical parameter). We can go a step further, and show that the second partial derivative of the entropy is negative at all points on the manifold of states for a system in the state space. We consider a similar situation to the above, where we have an isolated system containing a gas. We then partition the system into two subsystems using some internal membrane and consider the parameters of the subsystems and total system. This state of the system is sometimes known as **partial equilibrium**, in which every subsystem is locally in equilibrium, while the composite system as a whole is not. The initial entropies of the subsystems, with the membrane separating them, are $S_A(E_A, \{X_A\})$ and $S_B(E_B, \{X_B\})$, while the total entropy at this point is simply the sum of the two subsystems' entropies. If we now remove the partition between the subsystems, we expect some equilibration process to occur which gives us a new equilibrium entropy $S_{\text{tot}}(E_{\text{tot}}, \{X_{\text{tot}}\})$. The second law tells us that we must have $S_{\text{tot}} \geq S_A + S_B$ (if we apply the free expansion idea to each subsystem individually, ignoring the gas filling the other part of the system). Suppose this partition is placed such that the energy E and other extensive parameters $\{X\}$ are split between the subsystems according to some parameter λ , that is:

$$E_{\text{tot}} = \lambda E_A + (1 - \lambda) E_B, \quad X_{\text{tot}}^i = \lambda X_A^i + (1 - \lambda) X_B^i$$

We can then write an inequality for the total entropy:

$$\begin{aligned} S_{\text{tot}} &= S(\lambda E_A + (1 - \lambda) E_B, \lambda X_A^i + (1 - \lambda) X_B^i) \\ &\geq S_A(\lambda E_A, \lambda X_A^i) + S_B((1 - \lambda) E_B, (1 - \lambda) X_B^i) \\ &= \lambda S_A(E_A, X_A^i) + (1 - \lambda) S_B(E_B, X_B^i) \end{aligned} \quad (2.10)$$

where we use the idea of extensivity in the last line. This is the defining relation for a strictly concave function, and establishes that the entropy S of an isolated system is a concave function of its extensive parameters. These findings are summarised in a green box here, so the folks who rapidly scroll through will get a chance to appreciate that this is a *very important* point.

Definition 2.4: Principle of Entropy Maximisation

The equilibrium value of any unconstrained extensive parameter is that which maximises the entropy at a fixed value of the internal energy. This implies that the entropy of a system is a concave function of its extensive parameters, and its maximum point determines the equilibrium values of its extensive parameters.

2.8 The Principle of Energy Minimisation

We can now show that the entropy maximisation principle naturally leads to the requirement that the internal energy of a system in equilibrium is minimised. This can be done in both a physical and mathematical capacity, and I will attempt to show you both for maximal clarity. We will start with the physical argument first.

Suppose we have a system that is in equilibrium, but is in a state which does not have the minimum internal energy. We can adiabatically extract this energy through work, lowering its internal energy. Then, we convert this work entirely into heat using an engine with zero efficiency and direct that heat back into the system. This brings the internal energy of the system back to its initial value, but the introduction of extra heat would have now increased the entropy of the system past its previous equilibrium value. This is in direct contradiction with the entropy maximisation principle! We thus conclude that the internal energy of the system must have been minimised at equilibrium, such that no useful work can be extracted from a system in an equilibrium state.

Now, we turn to the mathematical argument which will aid our geometric interpretation of thermodynamics. We first consider the first derivative at equilibrium (which we expect to vanish), denoting this quantity by Y :

$$Y \equiv \left(\frac{\partial E}{\partial X} \right)_S = - \frac{(\partial S / \partial X)_E}{(\partial S / \partial E)_X} = -T \left(\frac{\partial S}{\partial X} \right)_E = 0 \quad (2.11)$$

where we have used the triple product rule in the first equality, and the thermodynamic definition of temperature¹⁷ in the second equality. Even without assuming that this is an extremal point, we see that the first partial derivative of the internal energy with respect to the entropy must vanish at equilibrium, as imposed by the entropy maximisation principle. We now look at the second derivative to determine if this extremum is a minimum or maximum point:

$$\begin{aligned} \left(\frac{\partial^2 E}{\partial X^2} \right)_S &= \left(\frac{\partial Y}{\partial X} \right)_S = \left(\frac{\partial Y}{\partial E} \right)_X \left(\frac{\partial E}{\partial X} \right)_S + \left(\frac{\partial Y}{\partial X} \right)_E \\ &= \underbrace{\left(\frac{\partial Y}{\partial E} \right)_X}_Y + \left(\frac{\partial Y}{\partial X} \right)_E \\ &= 0 \text{ at equilibrium} \\ &= - \frac{\partial}{\partial X} \left[\frac{(\partial S / \partial X)_E}{(\partial S / \partial E)_X} \right]_E \\ &= - \frac{\partial^2 S / \partial X^2}{\partial S / \partial E} + \frac{\partial S}{\partial X} \frac{\partial^2 S / \partial X \partial E}{(\partial S / \partial E)^2} \\ &= -T \frac{\partial^2 S}{\partial X^2} > 0 \quad \text{at} \quad \frac{\partial S}{\partial X} = 0 \end{aligned} \quad (2.12)$$

which confirms that the internal energy is minimised at equilibrium, agreeing with the above. In a very similar fashion to the above, we can also show that the internal energy is a convex function of its extensive parameters. Once again, we will box this up in green to ensure that everyone has a fair chance of seeing this content.

¹⁷Technically I only define this in a later section, but having established that $\delta Q = T dS$, this should be clear by this point.

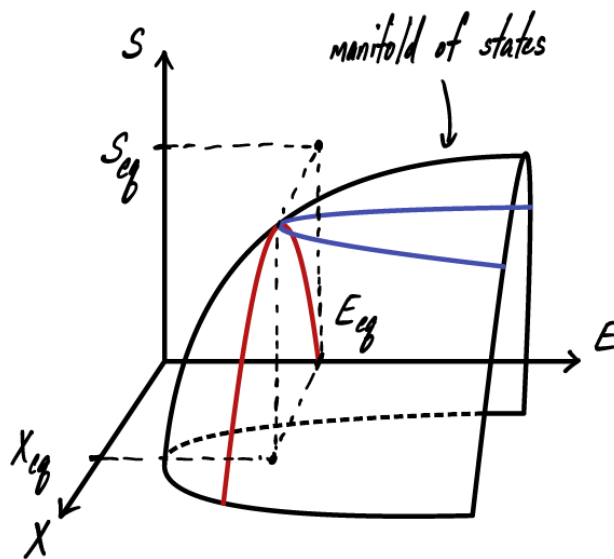


FIGURE 10: Entropy Maximisation and Energy Minimisation on a Manifold. The manifold of states available to a system in the space of extensive parameters must take on a form such that the entropy is concave in the other parameters, while the internal energy is convex in the other parameters. This gives maxima for the curves of fixed internal energy (red), which determine the equilibrium parameters of the system and also coincide with the minima of the curves of fixed entropy (blue).

Definition 2.5: Principle of Energy Minimisation

The equilibrium value of any unconstrained extensive parameter is that which minimises the internal energy at a fixed value of the entropy. This implies that the internal energy of a system is a convex function of its extensive parameters, and its minimum point determines the equilibrium values of its extensive parameters.

Geometrically, we can visualise the structure of a manifold of states available to a system as being a surface function S as a function of the internal energy E and other extensive parameters $\{X\}$. This function must be single-valued and concave, which then naturally leads to a picture of convexity when rotated to view the surface as that of the energy E in terms of the entropy S and extensive parameters $\{X\}$. This is the beginning of the geometric picture of thermodynamics, as constructed by Constantin Carathéodory in 1909. This leads to a very rigorous (yet equivalent) framework of thermodynamics, rooted in differential geometry and partial differential equations while still agreeing with the predictions of statistical mechanics. While acknowledging that this picture exists (and I strongly encourage you to read up on it if you are interested), we will typically only make references to the ideas of Carathéodory's geometric picture, without rigorously developing the ideas in the same way that he did.

2.9 Total Differentials and Measurability

Having established the existence of a path-independent state variable S that characterises the heat exchanged by a system, we now want to study how this plays into the first law. We first assume that a particular system takes on an equation of state of the form $f(p, V, T) = 0$, and consider the total differentials of the energy E and entropy S in terms of measurable quantities such as V, T, C_V , etc. Specifically, we want to show that finite

changes in the energy and entropy of such a system can be measured without any knowledge of the equation of state.

Since there exists an equation of state constraining one of $\{p, V, T\}$, we can choose one variable to be dependent — we choose the pressure p for this. Then, we begin with the total differential of the internal energy:

$$dE = \left(\frac{\partial E}{\partial T}\right)_V dT + \left(\frac{\partial E}{\partial V}\right)_T dV$$

From the definition of the entropy as a state variables, we write:

$$\begin{aligned} dS &= \frac{dQ}{T} = \frac{dE}{T} + \frac{p dV}{T} \\ &= \frac{1}{T} \left(\frac{\partial E}{\partial T}\right)_V dT + \frac{1}{T} \left[\left(\frac{\partial E}{\partial V}\right)_T + p \right] dV \\ &\equiv \left(\frac{\partial S}{\partial T}\right)_V dT + \left(\frac{\partial S}{\partial V}\right)_T dV \end{aligned}$$

where we write the last line with the knowledge the S has an exact differential. Using Schwarz' theorem, we now compare the second derivatives of S by taking the cross-partial derivatives of each partial derivative:

$$\left[\frac{\partial}{\partial V} \left(\frac{\partial S}{\partial T}\right)_V \right]_T = \left[\frac{\partial}{\partial T} \left(\frac{\partial S}{\partial V}\right)_T \right]_V$$

Thus, we can enact these cross-partial derivatives on the above expression for dS to obtain:

$$\frac{1}{T} \left[\frac{\partial}{\partial V} \left(\frac{\partial E}{\partial T}\right)_V \right]_T = -\frac{1}{T^2} \left(\frac{\partial E}{\partial V}\right)_T + \frac{1}{T} \left[\frac{\partial}{\partial T} \left(\frac{\partial E}{\partial V}\right)_T \right]_V - \frac{p}{T^2} + \frac{1}{T} \left(\frac{\partial p}{\partial T}\right)_V$$

with the cancellation performed using Schwarz' theorem again. Finally, multiplying through by T^2 gives us:

$$p + \left(\frac{\partial E}{\partial V}\right)_T - T \left(\frac{\partial p}{\partial V}\right)_V = 0 \quad (2.13)$$

Returning to the total differential of the internal energy, we see that $(\partial E/\partial V)_T$ can be eliminated using Eq. 2.13, while $(\partial E/\partial T)_V$ is simply the constant-volume heat capacity C_V . This returns:

$$dE = C_V dT + \left[T \left(\frac{\partial p}{\partial T}\right)_V - p \right] dV \quad (2.14)$$

Returning to the total differential of the entropy, we see that the prefactor of dV can be eliminated using Eq. 2.13 to obtain:

$$dS = \frac{C_V}{T} dT + \left(\frac{\partial p}{\partial T}\right)_V dV \quad (2.15)$$

Thus, the two total differentials in Eqs. 2.14 and 2.15 are now solely in terms of measurable quantities. Each of these can be measured as follows:

- T : Use a thermometer.
- p : Use a manometer.
- C_V : Use a calorimeter.

- $(\partial p/\partial T)_V$: Use a manometer and thermometer with a box.

Having dealt with the total differential of entropy, we see that the heat capacity naturally appears in its expression. This can be formalised and generalised to other heat capacities for some other constant state variables. From Eq. 2.15, we see that we can hold the volume of our system constant to obtain:

$$C_V = T \left(\frac{\partial S}{\partial T} \right)_V \quad (2.16)$$

We can obtain a similar expression for C_p , albeit with a little more work. From Eq. 2.15 again, we convert the dependencies of V to that of p by using the total differential of $V(p, T)$:

$$\begin{aligned} dS &= \frac{C_V}{T} dT + \left(\frac{\partial p}{\partial T} \right)_V dV \\ &= \frac{C_V}{T} dT + \left(\frac{\partial p}{\partial T} \right)_V \left[\left(\frac{\partial V}{\partial T} \right)_p dT + \left(\frac{\partial V}{\partial p} \right)_T dp \right] \\ &= \underbrace{\left[\frac{C_V}{T} + \left(\frac{\partial p}{\partial T} \right)_V \left(\frac{\partial V}{\partial T} \right)_p \right]}_{\left(\frac{\partial S}{\partial T} \right)_p} dT + \underbrace{\left(\frac{\partial p}{\partial T} \right)_V \left(\frac{\partial V}{\partial p} \right)_T}_{\left(\frac{\partial S}{\partial p} \right)_T} dp \end{aligned}$$

where we identify in the last line that the coefficients of each individual differential are equal to the partial derivatives of S using the total differential dS . We now look at the coefficient of dT :

$$\begin{aligned} \left(\frac{\partial S}{\partial T} \right)_p &= \frac{C_V}{T} + \left(\frac{\partial p}{\partial T} \right)_V \left(\frac{\partial V}{\partial T} \right)_p \\ T \left(\frac{\partial S}{\partial T} \right)_p &= C_V + T \left(\frac{\partial p}{\partial T} \right)_V \left(\frac{\partial V}{\partial T} \right)_p \end{aligned}$$

However, the second term on the right can be modified using Eq. 2.13 to give:

$$T \left(\frac{\partial S}{\partial T} \right)_p = C_V + \left[p + \left(\frac{\partial E}{\partial V} \right)_T \right] \left(\frac{\partial V}{\partial T} \right)_p$$

This is nothing more than the result we had previously obtained in Eq. 1.11! Thus, we can write the constant-pressure heat capacity as:

$$C_p = T \left(\frac{\partial S}{\partial T} \right)_p \quad (2.17)$$

The resemblance of this expression for C_p to that of Eq. 2.16 is no coincidence, and is really just a specific case of a more general relationship that is easily seen. Knowing that the inexact differential for heat is now given by $\delta Q = T dS$, we can use the definition of the heat capacity from Eq. 1.6 to write:

$$\left(\frac{\delta Q}{\delta T} \right)_y = T \left(\frac{\partial S}{\partial T} \right)_y \equiv C_y \quad (2.18)$$

Note that this is not obtained by simply taking the partial derivative of $T dS$ with respect to T at constant y , but rather is obtained by dividing $T \Delta S$ by a finite ΔT , before taking the limit of $\Delta T \rightarrow 0$ at constant y .

As a final point, we can also consider rewriting the heat capacity difference using our new knowledge of the entropy. This is valuable in experimental contexts, since a measurement of C_p is significantly harder to perform than one of C_V . It turns out that we can rewrite the heat capacity difference purely in terms of

response functions, which are far easier to study in the lab. We start by considering the total differential of $S(T, V)$. Dividing this by finite ΔT and taking the $\Delta T \rightarrow 0$ limit at constant p , we obtain:

$$\begin{aligned}\left(\frac{\partial S}{\partial T}\right)_p &= \left(\frac{\partial S}{\partial T}\right)_V + \left(\frac{\partial S}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_p \\ \frac{C_p}{T} &= \frac{C_V}{T} + \left(\frac{\partial S}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_p \\ \implies C_p - C_V &= T \left(\frac{\partial S}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_p\end{aligned}$$

Now, we invoke the definition of the isobaric (constant p) to eliminate $(\partial V/\partial T)_p$. To get rid of the $(\partial S/\partial V)_T$ term, we use the result from Eq. 2.15, where we have $(\partial S/\partial V)_T = (\partial p/\partial T)_V$. This returns:

$$C_p - C_V = T \left(\frac{\partial p}{\partial T}\right)_V \left(\frac{\partial V}{\partial T}\right)_p$$

This nearly resembles the compressibility κ , except with a permutation of all the variables. We thus fix this by invoking the triple product rule:

$$\left(\frac{\partial X}{\partial Y}\right)_Z \left(\frac{\partial Y}{\partial Z}\right)_X \left(\frac{\partial Z}{\partial X}\right)_Y = -1$$

for any set of variables $\{X, Y, Z\}$ which are expressible in terms of each other. Applying this to $(\partial p/\partial T)_V$, we have:

$$\left(\frac{\partial p}{\partial T}\right)_V = - \left(\frac{\partial V}{\partial T}\right)_p \left(\frac{\partial V}{\partial p}\right)_T^{-1}$$

Finally, we plug all of this into the heat capacity difference to arrive at:

$$\begin{aligned}C_p - C_V &= -T \left(\frac{\partial V}{\partial T}\right)_p^2 \left(\frac{\partial V}{\partial p}\right)_T^{-1} \\ &= \frac{TV\beta_p}{\kappa_T}\end{aligned}\tag{2.19}$$

where β_p is the isobaric expansivity and κ_T is the isothermal compressibility. Thus, if one is able to measure C_V for a given system, C_p can be obtained by simply measuring two volume-response functions instead.

3 Thermodynamic Potentials

The typical problem statement in thermodynamics is one where we consider an initially isolated system with total energy E_0 , volume V_0 and particle number N_0 . These numbers are always conserved for the entire isolated system as a whole (otherwise it won't be isolated). Within the system, there are then two or more subsystems that are separated by internal partitions. These partitions must be adiabatic (preventing E transfer), rigid (preventing V transfer) and impermeable (preventing N transfer). Thus, to set the problem in motion, one or more of these partitions must be relaxed to allow some form of transfer. This sets off spontaneous processes, which ultimately allow the system to settle into a new equilibrium state. In determining the new equilibrium values of each partitioned region, we want to get rid of all forms of path dependence in the independent variables, so that knowing the initial and final states is sufficient to completely determine the changes in $E^{(i)}$, $V^{(i)}$ and $N^{(i)}$ for the i -th partitioned region. For this, we must return to the first law of thermodynamics.

3.1 The Fundamental Thermodynamic Relation

The first and second laws of thermodynamics can be combined to produce what many call the **fundamental thermodynamic relation** in energy representation:

$$dE = T dS + \underbrace{\sum_{i=1}^r F_i dx_i}_{\text{mechanical}} + \underbrace{\sum_{j=1}^s \mu_j dN_j}_{\text{chemical}} \quad (3.1)$$

The first term on the right hand side represents the heat change, while the next two terms represent all possible forms of work that may be done on the system. Broadly, these are classified into either mechanical work or chemical work. For mechanical work, the extensive variable being changed is typically some sort of linear dimension, with a conjugate intensive variable corresponding to the constant quantity at equilibrium (e.g. allowing for a system to exchange volume with its surroundings ensures that it is isobaric with the environment). For chemical work, the extensive variable is typically a particle number, with the conjugate intensive variable being an associated chemical potential that governs the ease of particle flow. We see that we have $(1 + r + s)$ degrees of freedom for the energy to change through the extensive variables, so we should also have $(1 + r + s)$ equations of state. These are simply obtained by writing the total differential of the internal energy as:

$$dE = \left(\frac{\partial E}{\partial S} \right)_{\{x_i\}, \{N_j\}} dS + \sum_{k=1}^r \left(\frac{\partial E}{\partial x_k} \right)_{S, \{x_{i \neq k}\}, \{N_j\}} dx_k + \sum_{k=1}^s \left(\frac{\partial E}{\partial N_k} \right)_{S, \{x_i\}, \{N_{j \neq k}\}} dN_k \quad (3.2)$$

Comparing this with Eq. 3.1, we thus obtain the $(1 + r + s)$ equations of state:

$$T = T(S, \{x_i\}, \{N_j\}) = \left(\frac{\partial E}{\partial S} \right)_{\{x_i\}, \{N_j\}} \quad (3.3)$$

$$F_k = F_k(S, \{x_i\}, \{N_j\}) = \left(\frac{\partial E}{\partial x_k} \right)_{S, \{x_{i \neq k}\}, \{N_j\}} \quad (3.4)$$

$$\mu_k = \mu_k(S, \{x_i\}, \{N_j\}) = \left(\frac{\partial E}{\partial N_k} \right)_{S, \{x_i\}, \{N_{j \neq k}\}} \quad (3.5)$$

Thus, knowledge of all the equations of state (as functions of the extensive variables only) gives us total knowledge of the total differential for the internal energy, which completely describes the system since it returns the fundamental thermodynamic relation.

3.2 The Euler Equation

In the energy representation for the fundamental relation, we see that E is an extensive function of $(1 + r + s)$ extensive variables. Thus, we can write:

$$E(\lambda S, \{\lambda x_i\}, \{\lambda N_i\}) = \lambda E(S, \{x_i\}, \{N_i\})$$

We now differentiate both sides of the above relation by the scaling parameter λ (take the total derivative $d/d\lambda$, not the partial derivative $\partial/\partial\lambda$):

$$\frac{\partial E(\lambda S, \{\lambda x_i\}, \{\lambda N_i\})}{\partial(\lambda S)} \frac{\partial(\lambda S)}{\partial \lambda} + \sum_{i=1}^r \frac{\partial E(\lambda S, \{\lambda x_i\}, \{\lambda N_i\})}{\partial(\lambda x_i)} \frac{\partial(\lambda x_i)}{\partial \lambda} + \sum_{j=1}^s \frac{\partial E(\lambda S, \{\lambda x_i\}, \{\lambda N_i\})}{\partial(\lambda N_j)} \frac{\partial(\lambda N_j)}{\partial \lambda} = E(S, \{x_i\}, \{N_i\})$$

Then, we simply set $\lambda = 1$ and use the equations of state from above to arrive at:

$$TS + \sum_{i=1}^r x_i F_i + \sum_{j=1}^s N_j \mu_j = E(S, \{x_i\}, \{N_i\}) \quad (3.6)$$

This result is known as the **Euler equation**, and is an incredibly useful relation which tells us that the energy of a system in its natural variables (S, V, N) can be expressible in terms of the products of all conjugate variable pairs, when they are also expressed in terms of the natural variables (S, V, N).

3.3 The Gibbs–Duhem Relation

The Euler equation provides us with a link between a thermodynamic quantity, such as the internal energy E , and its intensive and extensive constituents. This also encodes another relationship which will reveal that intensive quantities are fundamentally constrained as compared to their extensive counterparts. To derive this, we first take the total differential of the Euler equation:

$$dE = T dS + S dT + \sum_{i=1}^r (F_i dx_i + x_i dF_i) + \sum_{j=1}^s (\mu_j dN_j + N_j d\mu_j)$$

We now use the fundamental relation from Eq. 3.1 to cancel corresponding terms from both sides, arriving at:

$$S dT + \sum_{i=1}^r x_i dF_i + \sum_{j=1}^s N_j d\mu_j = 0 \quad (3.7)$$

This is known as the **Gibbs–Duhem relation**, and it tells us that there are only $(r + s)$ independent intensive variables (as compared to $(1 + r + s)$ independent extensive variables) in describing a thermodynamic system. That is, one intensive variable can always be expressed in terms of all others. This relation hints at a more fundamental geometric theory underlying thermodynamics, since the dimensionality of a purely intensive subspace is smaller than an extensive one. We will discuss these implications in a later section.

3.4 The Entropy Representation

In a similar manner to the energy representation for the fundamental relation and its associated thermodynamics, we also have the entropy representation, where the total differential of the entropy is instead used to derive equations of state in terms of its natural extensive variables (E, V, N). This is written:

$$dS = \left(\frac{\partial S}{\partial E} \right)_{\{x_i\}, \{N_j\}} dE + \sum_{i=1}^r \left(\frac{\partial S}{\partial x_i} \right)_{E, \{x_{k \neq i}\}, \{N_j\}} dx_i + \sum_{j=1}^s \left(\frac{\partial S}{\partial N_j} \right)_{E, \{x_i\}, \{N_{k \neq j}\}} dN_j \quad (3.8)$$

Comparing the partial derivatives to Eq. 3.1, we thus obtain the equations of state for the entropy representation as:

$$\frac{1}{T} = \frac{1}{T(E, \{x_i\}, \{N_j\})} = \left(\frac{\partial S}{\partial E} \right)_{\{x_i\}, \{N_j\}} \quad (3.9)$$

$$\frac{F_i}{T} = \frac{F_i(S, \{x_i\}, \{N_j\})}{T} = - \left(\frac{\partial S}{\partial x_i} \right)_{E, \{x_{k \neq i}\}, \{N_j\}} \quad (3.10)$$

$$\frac{\mu_j}{T} = \frac{\mu_j(S, \{x_i\}, \{N_j\})}{T} = - \left(\frac{\partial S}{\partial N_j} \right)_{S, \{x_i\}, \{N_{k \neq j}\}} \quad (3.11)$$

which returns the fundamental relation in a more useful form as:

$$dS = \frac{1}{T} dE - \sum_{i=1}^r \frac{F_i}{T} dx_i - \sum_{j=1}^s \frac{\mu_j}{T} dN_j \quad (3.12)$$

This breaks down the entropy as a function of $(1+r+s)$ extensive variables, with corresponding conjugate intensive variables as the coefficients. Since this is parametrised by the same number of extensive quantities as the internal energy, we say that the energy and entropy representations are equivalent.

3.5 Conditions for Thermodynamic Equilibrium

To study the conditions required for thermodynamic equilibrium, we can use our newly developed idea of the second law to impose bounds on how a system's entropy should behave. For simplicity, we will only consider mechanical work in the form of $(-p dV)$, along with a single particle species (so $s = 1$). We first consider a system partitioned into two subsystems, such that the total system is isolated and out of equilibrium to begin with. The qualifier that the system is isolated now implies that the total energy E , total volume V and total particle number N are conserved. Then, we notice that the entropy being an extensive function implies that it must be additive for the two systems. Mathematically, we say:

$$S(E, V, N) = S^{(1)}(E^{(1)}, V^{(1)}, N^{(1)}) + S^{(2)}(E^{(2)}, V^{(2)}, N^{(2)})$$

Since the system is isolated, we know that $dS \geq 0$, with equality only achieved either when the system is undergoing a reversible process, or when thermodynamic equilibrium has already been attained (and hence no state variables are changing). We can then consider the total differential for the entropy of the isolated system:

$$\begin{aligned} dS &= dS^{(1)} + dS^{(2)} \\ &= \left(\frac{\partial S^{(1)}}{\partial E^{(1)}} \right)_{V^{(1)}, N^{(1)}} dE^{(1)} + \left(\frac{\partial S^{(1)}}{\partial V^{(1)}} \right)_{E^{(1)}, N^{(1)}} dV^{(1)} + \left(\frac{\partial S^{(1)}}{\partial N^{(1)}} \right)_{E^{(1)}, V^{(1)}} dN^{(1)} \\ &\quad + \left(\frac{\partial S^{(2)}}{\partial E^{(2)}} \right)_{V^{(2)}, N^{(2)}} dE^{(2)} + \left(\frac{\partial S^{(2)}}{\partial V^{(2)}} \right)_{E^{(2)}, N^{(2)}} dV^{(2)} + \left(\frac{\partial S^{(2)}}{\partial N^{(2)}} \right)_{E^{(2)}, V^{(2)}} dN^{(2)} \\ &= \left(\frac{1}{T^{(1)}} - \frac{1}{T^{(2)}} \right) dE^{(1)} + \left(\frac{p^{(1)}}{T^{(1)}} - \frac{p^{(2)}}{T^{(2)}} \right) dV^{(1)} - \left(\frac{\mu^{(1)}}{T^{(1)}} - \frac{\mu^{(2)}}{T^{(2)}} \right) dN^{(1)} \end{aligned}$$

where in the last line, we use the fact that $dE^{(1)} + dE^{(2)} = dE = 0$, and likewise for the other extensive variables. Now, setting $dS = 0$ at equilibrium tells us that each term on the right hand side must have its coefficient vanish. This leads us to the three conditions of thermodynamic equilibrium on the intensive variables:

$$T^{(1)} = T^{(2)} \quad (3.13)$$

$$p^{(1)} = p^{(2)} \quad (3.14)$$

$$\mu^{(1)} = \mu^{(2)} \quad (3.15)$$

To better understand how these intensive variables characterise a system *approaching* thermodynamic equilibrium, we consider the same setup but now relax the condition that $dS = 0$. For simplicity, we will assume that the pressures and chemical potentials are equal, and only a temperature differential exists, since the argument for the other two cases is identical. In the approach to equilibrium, each of the subsystems acts as a heat bath for the other, with the two having a finite temperature difference. We aim to show that a temperature differential between the subsystems will spontaneously induce a flow of energy to lower this differential. For this, we consider the system approaching equilibrium at one instant of time:

$$dS > 0 \implies \left(\frac{1}{T^{(1)}} - \frac{1}{T^{(2)}} \right) dE^{(1)} > 0$$

In the next instant, the energy of each subsystem would have changed due to this relation. Without loss of generality (WLOG), we assume:

$$E^{(1)} \longrightarrow E^{(1)} + dE^{(1)}, \quad E^{(2)} \longrightarrow E^{(2)} - dE^{(1)}$$

Now, if $T^{(1)} > T^{(2)}$, then we see that $dE^{(1)} < 0$, implying that energy flows from subsystem 1 into subsystem 2. Likewise, if $T^{(2)} > T^{(1)}$, then we see that $dE^{(1)} > 0$, implying that energy flows from subsystem 2 into subsystem 1 instead. This demonstrates that any isolated system undergoing a process of equilibration will spontaneously transfer energy, volume and particles in order to reach its equilibrium set of extensive variables, in perfect agreement with Def. 0.1.

3.6 Legendre Transformations

In real thermodynamic problems, it is very rare that we will be able to truly keep constant the variables shown in the subscripts of the partial derivatives. For instance, the addition of particles ($dN > 0$) at constant entropy¹⁸ seems like a wildly impossible thing to do in a real experimental setting, yet this is the sort of precise control the equations of state require when working in the energy representation. We thus need to find a way to modify the fundamental relation to suit physical problems where the variables within our control are not all extensive variables, and this is where we have to introduce a neat mathematical tool.

We first consider some function $f(x)$ whose independent variable is x . In the space of \mathbb{R}^2 , the function is specified as the locus of points $(x, f(x))$, implying that two parametrised degrees of freedom (equivalently, coordinates) are required to construct said locus. Suppose we now wanted to instead construct the locus using the gradient (conjugate variable to x) of the function $\partial f / \partial x = \alpha$ as the independent variable, rather than the x values. The naïve guess here might be to simply obtain an expression of the form:

$$\frac{\partial f}{\partial x} = \alpha(x) \implies x = x(\alpha)$$

by inverting the relationship of the gradient in terms of x . We could then simply plug this back into $f(x)$ to obtain a new function $f(x(\alpha)) \equiv g(\alpha)$. However, this would constitute a massive loss of information, as there are infinitely many curves which would obey the same relation up to a translation parallel to the x -axis. In the

¹⁸Technically, this could be done as an adiabatic process, though it is still hard.

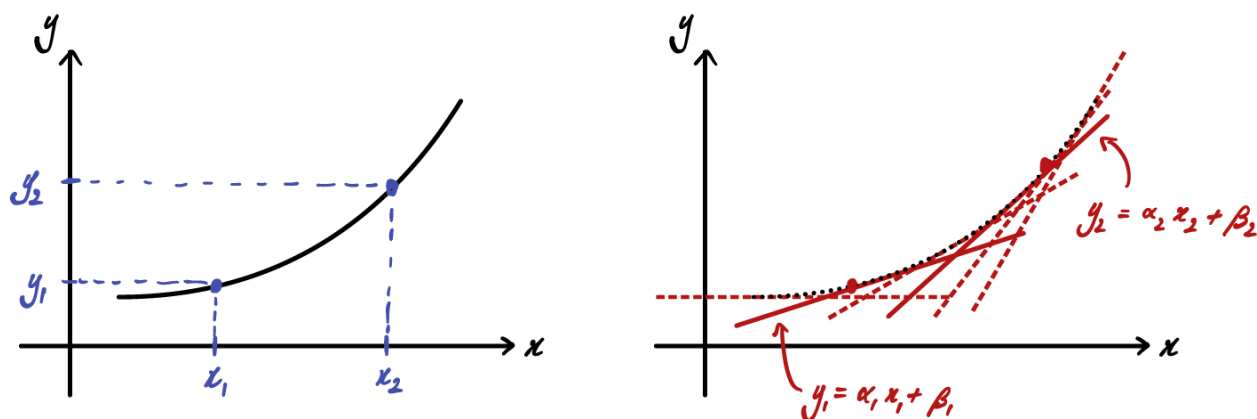


FIGURE 11: Duality between Point Geometry and Line Geometry. (Left) A curve is constructed as a locus of points satisfying some condition $(x, f(x))$. Each point is uniquely specified by two degrees of freedom – the point coordinates. (Right) Equivalently, a locus can be constructed using a family of lines tangent to the curve at each point, with the two degrees of freedom now being given in the form of the slopes and intercepts.

context of thermodynamics, this would be equivalent to an attempt at expressing the fundamental relation in terms of the temperature rather than the entropy, but would be an incomplete description of the system owing to the mathematical loss of information. We thus need another degree of freedom to rectify this inadequacy.

Instead of specifying a locus by its parametrised set of coordinates, one can also specify it using a family of lines which are tangent to the curve at each point. These tangent lines also have two degrees of freedom: the slope and y -intercept. By specifying a set of slopes α and the corresponding intercepts β , we obtain a set of fixed lines whose envelope¹⁹ then reproduces the original locus. This allows us to constrain the gradient values via a new function $\beta(\alpha)$, which takes in the gradient as the independent value and returns the y -intercept. For a given point $(x, f(x))$, this is not too hard to compute:

$$\alpha = \frac{f(x) - \beta}{x} \implies \beta(\alpha) = f(x) - \alpha x$$

This process of obtaining the new function $\beta(\alpha)$ from $f(x)$ is known as the **Legendre transformation**. While we have only considered a function of a single variable here, the process is easily generalised to multivariate functions, since the partial derivative keeps all other independent variables constant, so Legendre transformations can be successively performed on each independent variable as required. In terms of its total differential, we have:

$$df = \left(\frac{\partial f}{\partial x} \right)_y dx + \left(\frac{\partial f}{\partial y} \right)_x dy \equiv u dx + v dy \quad (3.16)$$

where $u = u(x, y) = (\partial f / \partial x)_y$ and $v = v(x, y) = (\partial f / \partial y)_x$ are simply the partial derivatives with respect to the constituent variables. The pairs of variables (u, x) and (v, y) are called **conjugate variable pairs**. Using the Legendre transformation, we can swap the dependency of a function between one of its variables and that corresponding variable's conjugate as:

$$g(u, y) = f(x(u, y), y) - x(u, y)u \quad (3.17)$$

¹⁹There is a lot of analytical work that goes into this definition of an *envelope of a set of lines*, but we will not let that prevent us from doing physics here. This can be studied in any elementary textbook on differential geometry, or even several advanced textbooks on classical dynamics.

where we assume that it is possible to invert $u(x, y)$ to obtain an expression for $x(u, y)$. Now, we check the total differential of this new function:

$$\begin{aligned} dg &= df - x du - u dx \\ &= u dx + v dy - x du - u dx \\ &= -x du + v dy \end{aligned}$$

In this, we see that the total differential of $g(u, y)$ has a pure dependency on u and y only! Furthermore, the conjugate variable swap on (u, x) was successfully performed, with x now taking on the role of $-(\partial g/\partial u)_y$. Thus, given a function $f(x, y)$, we can describe the general procedure for enacting a Legendre transformation on the conjugate variable pair (u, x) as follows:

Definition 3.1: Procedure for Legendre Transformations

To transform a function $f(x, y)$ into a new function $g(u, y)$, where the variable x is expressible in terms of $x(u, y)$, we use the following procedure:

1. Define $g(u, y) = f(x, y) - xu$ as the new transformed function.
2. For every appearance of x in g , replace this with $x = x(u, y)$.
3. Rewrite the total differential of g using the new variables u and y .

3.7 The Four Fundamental Thermodynamic Potentials

We will now explore how the Legendre transformation is used in turning thermodynamic potentials (such as the internal energy) into other potentials which are more useful for suitably controlled systems. For this section, I will assume mechanical work only comes in the form of $p dV$ and chemical work is only done on a single species (this just implies that $s = 1$)²⁰. We begin with the fundamental thermodynamic relation in the energy representation:

$$dE = T dS - p dV + \mu dN \quad (3.18)$$

This form of the internal energy is our first thermodynamic potential. It should be noted that this is only true when it is expressed in terms of the extensive variables S, V and N , otherwise it loses a great deal of meaning in thermodynamics. We now transform E into another potential by enacting the Legendre transformation on the (T, S) conjugate variable pair, obtaining:

$$F(T, V, N) = E(S(T, V, N), V, N) - TS(T, V, N) \quad (3.19)$$

$$dF = -S dT - p dV + \mu dN \quad (3.20)$$

This new potential $F(T, V, N)$ is called the **Helmholtz free energy**, and has the canonical variables of T, V and N . It is the suitable potential to be used when a system is allowed to exchange heat with its environment through coupling with a large reservoir, enabling equilibration to a fixed temperature T . It should also be

²⁰The generalisation is quite simple, and will only make the notation look clunkier here with little added intuition (law of diminishing returns).

noted that in Eq. 3.19, we should really be using $\langle E \rangle$ instead of E , to explicitly denote that we are interested in the macroscopic energy, and not the energy of any one given microstate. Another potential is obtained by taking the Legendre transform of the internal energy, only this time we enact it on the conjugate variable pair (p, V) to obtain:

$$H(S, p, N) = E(S, V(S, p, N), N) + pV(S, p, N) \quad (3.21)$$

$$dH = T dS + V dp + \mu dN \quad (3.22)$$

This new potential $H(S, p, N)$ is called the **enthalpy**, and has the canonical variables of S, p and N . The enthalpy is the ideal potential for systems which are allowed to exchange volume with their environment but not heat, reaching an equilibrium at a fixed pressure p . Finally, we can take the Legendre transform of the (T, S) conjugate variable pair on the enthalpy (or equivalently of the (p, V) conjugate variable pair on the Helmholtz free energy) to obtain:

$$G(T, p, N) = E - TS + pV \quad (3.23)$$

$$dG = -S dT + V dp + \mu dN \quad (3.24)$$

This final potential $G(T, p, N)$ is called the **Gibbs free energy**, and has the canonical variables of T, p and N . As you may have guessed, this potential is ideal for systems which exchange both heat and volume with their surroundings, thus enabling an equilibrium characterisation at fixed temperature and pressure. At this point, one may ask why we can't simply Legendre transform the final conjugate variable pair (μ, N) as well. Indeed, it is possible to do so, but this would give us a "potential" that is fully defined in terms of intensive variables. As we established previously, there is always one less degree of freedom in a set of conjugate intensive variables as there are present in the extensive variables, so this final potential turns out to identically vanish since its total differential is exactly given by the Gibbs–Duhem relation.

3.8 Maxwell Relations

To wrap up this section, we will explore a neat consequence of the partial derivatives involved in the equations of state. Thusfar, we have seen that all thermodynamic equations of state are essentially just partial derivatives of the form $(\partial X / \partial Y)_{Z, W}$. Despite there being many such derivatives, the extent of their independence is limited by Schwarz' theorem, which links the second partial derivatives by symmetry. Specifically, Schwarz' theorem states:

$$\frac{\partial}{\partial Y'} \left[\left(\frac{\partial X}{\partial Y} \right)_{Z, Z'} \right] = \frac{\partial}{\partial Y} \left[\left(\frac{\partial X}{\partial Y'} \right)_{Z', Z} \right] \quad (3.25)$$

which is simply a statement that partial derivatives commute. Given a thermodynamic potential in its $(1 + r + s)$ natural variables, there are then $\binom{1+r+s}{2}$ unique pairs of mixed second partial derivatives. Evaluating the inner partial derivative in each of these pairs then links the first partial derivatives of two equations of state, constituting what is known as a **Maxwell relation**. For any thermodynamic potential of $(1 + r + s)$ variables, there are thus $\binom{1+r+s}{2}$ Maxwell relations available. As an example, we consider the available Maxwell relations for the internal energy (assuming $r = s = 1$). From Eq. 3.18, we take all possible mixed second partial derivatives and impose equality via Schwarz' theorem to write:

$$\left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial p}{\partial S}\right)_V = \frac{\partial^2 E}{\partial S \partial V} \quad (3.26)$$

$$\left(\frac{\partial T}{\partial N}\right)_S = \left(\frac{\partial \mu}{\partial S}\right)_N = \frac{\partial^2 E}{\partial S \partial N} \quad (3.27)$$

$$-\left(\frac{\partial p}{\partial N}\right)_V = \left(\frac{\partial \mu}{\partial V}\right)_N = \frac{\partial^2 E}{\partial V \partial N} \quad (3.28)$$

The Maxwell relations for all other thermodynamic potentials can be derived in an identical fashion.

3.9 Minimisation of the Helmholtz Free Energy

Previously, we had established extremal principles for the two thermodynamic potentials of entropy and internal energy in isolated systems. Specifically, we showed that the internal energy and entropy obey:

$$dS = 0, \quad d^2 S < 0 \quad (3.29)$$

$$dE = 0, \quad d^2 E > 0 \quad (3.30)$$

where these differentials refer to infinitesimal changes in the potentials (as a result of infinitesimal changes in the underlying extensive variables) which take the system away from equilibrium²¹. We now want to generalise these extremal principles to the new potentials obtained from the Legendre transformations, such that we can study the extremal properties of systems coupled to some reservoir. I will denote the entropy, energy and other extensive variables of the small closed subsystem by (S, E, \dots) and those of the reservoir by (S_R, E_R, \dots) , assuming the total system containing the subsystem and reservoir form an isolated system. We now apply the extremal principles to the combined system:

$$d(S + S_R) = 0, \quad d^2(S + S_R) < 0$$

$$d(E + E_R) = 0, \quad d^2(E + E_R) > 0$$

Now, we consider the reservoir's physical properties as opposed to those of the subsystem and demand that it is incredibly large in comparison. This means that it will realistically have an effectively infinite heat capacity C_R . Physically, this arises from the intuition that transferring any finite amount of heat in or out of the subsystem will have no measurable effect on the reservoir's temperature. Accordingly, the second differentials will behave as:

$$\frac{\partial^2 E_R}{\partial S_R^2} = \frac{\partial T_R}{\partial S_R} = \frac{T_R}{C_R} \longrightarrow 0$$

We thus conclude that the second-order changes in the bath quantities do not contribute to the total differentials, so we have:

$$d(S + S_R) = 0, \quad d^2 S < 0$$

$$d(E + E_R) = 0, \quad d^2 E > 0$$

Assuming that the subsystem and reservoir are only coupled through heat transfer and the combined total system is isolated, we can write the fundamental relation (in energy representation) for the reservoir as:

²¹I know that d^2 is a horrible abuse of notation, but you know exactly what I mean.

$$dE_R = T_R dS_R$$

Now, we use the first-order differentials of the entropy and internal energy to switch the reservoir quantities to those of the system:

$$dE = T_R dS \implies d(E - T_R S) = 0$$

with caution given to the signs on either side. Finally, we recognise that the condition of thermal equilibrium between the subsystem and reservoir demands that $T = T_R$ ²², so this differential simply states that the free energy $F = E - TS$ of the subsystem is extremised at equilibrium! The stationarity of the temperature then also allows us to write:

$$d^2(T_R S) = T_R d^2 S < 0$$

Together with $d^2 E > 0$, this gives us a positive second-order differential which specifies that the extremal point of the free energy is a minimal point! This is known as the principle of Helmholtz **free energy minimisation**. It should be noted that while this demonstrates the convexity of F in its extensive parameters, the Legendre transformation performed on the extensive parameter S now implies that F is a concave function of the intensive parameter T .

Definition 3.2: Principle of Free Energy Minimisation

The equilibrium value of any unconstrained internal parameter in a system that is in thermal contact with a heat reservoir is that which minimises the Helmholtz free energy of the system, within the manifold of states for which $T = T_R$. This implies that the Helmholtz free energy of a system is a convex function of its extensive parameters and a concave function of its intensive parameters.

4 The Third Law of Thermodynamics

With some basic intuition for the concept of entropy now established, we turn to study its behaviour at low temperatures, which reveals some interesting consequences for physical systems in general. From the definition of finite entropy changes:

$$\Delta S = S_f - S_i = \int_C \frac{dQ}{T}$$

we expect that there should be some special treatment to account for possibly singular behaviour as $T \rightarrow 0$. This motivates a third law of thermodynamics, to ensure that the entropy remains well-behaved at the absolute zero of the temperature scale while allowing for a definition of zero entropy that is consistent with statistical mechanics. This way, entropy can be treated in an absolute manner (as opposed to finite differences) in a similar manner to the absolute temperature. For systems that are coupled to volume or particle reservoirs, the minimisation principles for their respective thermodynamic potentials can be similarly shown.

²²This also implies that $dT = 0$ at equilibrium, so we can pull the temperature into the differential without consequence.

4.1 The Nernst Postulate

In 1907, Walther Nernst published a series of lectures titled *Experimental and Theoretical Applications of Thermodynamics to Chemistry*. In it, he noted a peculiar result where the Gibbs free energy and enthalpy of a system tended to the same constant value in the low temperature limit. With experimental verification at low temperatures, he continued to study this phenomenon for several years before putting forth his statement of the **third law of thermodynamics**, commonly known as the **Nernst postulate** [1].

Definition 4.1: Third Law of Thermodynamics

The Nernst postulate states that the entropy of a thermodynamic system goes to a constant for any set of thermodynamic variables $\{X_i\}$ as the temperature goes to zero. Mathematically, this is written as:

$$\lim_{T \rightarrow 0} S(T, \{X_i\}) = S_0 \quad (4.1)$$

This constant is independent of any of the thermodynamic variables, including the temperature.

Nernst's statement has shown to be true for all known real systems, including quantum mechanical systems. There is also a stronger statement by Max Planck, which states that the entropy of a thermodynamic system should tend towards zero as the temperature goes to zero. Planck's statement is true for most idealised classical systems, but fundamentally breaks down for quantum systems where the spin degree of freedom is still present. In the zero temperature limit, the presence of a spin degree of freedom (with no magnetic field, and hence a degeneracy in the spin states) allows the entropy to tend to $k_B \log 2$ per particle²³.

As an example of how the third law is applied to a real system in the form of the Nernst postulate, we can consider the entropy for a classical ideal gas based on the total differential we obtained in Eq. 2.15. For an ideal gas, we easily see that this can be integrated to give:

$$S(T, V) = C_V \log \left(\frac{T}{T_0} \right) + Nk_B \log \left(\frac{V}{V_0} \right) + S_0$$

where the integration constants appear as some reference temperature T_0 and volume V_0 , along with an arbitrary constant S_0 . In the zero temperature limit, we see that the entropy then tends to $-\infty$, which is a violation of the third law. This leads us to the conclusion that the classical ideal gas model is unphysical, and cannot possibly give meaningful results in the low temperature limit.

4.2 Vanishing Heat Capacities

We now explore some related consequences of the third law on the response functions for a physical system. For this, we first consider heating a system with all extensive thermodynamic variables kept constant. This gives the entropy change as:

$$\Delta S = S(T, \{X_i\}) - S(0, \{X_i\}) = \int_0^T dT' \frac{C_{\{X_i\}}(T')}{T'}$$

For a well-behaved function as $T \rightarrow 0$, we demand that this integral must converge. To deal with this more simply, we can expand the heat capacity as a power series in the temperature:

²³This is easily seen using the Gibbs entropy formula.

$$C_{\{X_i\}}(T) = \sum_{n=1}^{\infty} c_n T^n$$

where we ignore all powers $n \leq 0$ since they will clearly result in divergent behaviour. Now, taking the zero temperature limit, we have:

$$\lim_{T \rightarrow 0} C_{\{X_i\}}(T) = \lim_{T \rightarrow 0} \sum_{n=1}^{\infty} c_n T^n = 0 \quad (4.2)$$

Thus, we conclude that the heat capacity must vanish in the zero temperature limit! This means that an infinitesimal addition of heat at zero temperature will result in an infinite increase in temperature, which is clearly unphysical and thus is a first hint that zero temperature itself is not physically attainable.

4.3 Vanishing Thermal Expansivities

We can also perform a similar analysis of the thermal expansivity of a system in the zero temperature limit. We first define the generalised thermal expansivity for any generalised displacement:

$$\beta_{F_i} = \frac{1}{x_i} \left(\frac{\partial x_i}{\partial T} \right)_{F_i} \quad (4.3)$$

where F_i and x_i form a conjugate variable pair for mechanical work. Now, we use the Maxwell relation obtained from the total differential of the Gibbs free energy:

$$\left(\frac{\partial S}{\partial F_i} \right)_T = \left(\frac{\partial x_i}{\partial T} \right)_{F_i}$$

to write the generalised expansivity as:

$$\beta_{F_i} = \frac{1}{x_i} \left(\frac{\partial S}{\partial F_i} \right)_T$$

Since we know that the entropy tends to a constant in the zero temperature limit according to the Nernst postulate, the partial derivative on the right must identically vanish, leaving us with the result that the generalised expansivity itself vanishes in the zero temperature limit!

4.4 The Unattainability of Absolute Zero

We are now ready to confront perhaps the most peculiar consequence of the third law — the unattainability of absolute zero temperature²⁴. To approach this, we first consider cooling a system by varying some parameter $X_1 \rightarrow X_2$ to force a temperature change $T_1 \rightarrow T_2$. The entropy change at any temperature in such a process is then given by:

$$S(T, X) = S_0 + \int_0^T dT' \frac{C_X(X, T')}{T'}$$

Now, we assume (for sake of contradiction) that we *can* cool the system from $T_1 > 0$ to absolute zero adiabatically through some parameter change $X_1 \rightarrow X_2$. Then, we must have:

²⁴Note that there is nothing here which really prevents a system from attaining negative temperatures. You should be aware by now that this is absolutely possible, it's just not possible to reach negative temperatures by going *through* absolute zero. Rather, one has to go *through* infinite temperature.

$$\int_0^{T_1} dT' \frac{C_X(X, T')}{T'} = S(X_1, T) - S(X_2, T) = 0$$

since the process was adiabatic, and hence reversible. However, note that $C_X(T \neq 0, X)$ is strictly positive, so the integral itself can never be zero, forcing a contradiction! Thus, we conclude that there is no such process which exists to cool a system to absolute zero in a reversible fashion, and hence there cannot be any irreversible process which does so either (since the entropy change will never be zero for such a situation).

5 Open Systems

Up until this point, we have been studying relatively simple systems consisting of a single species typically contained in an isolated environment. In our consideration of the Helmholtz free energy, we also coupled the system to a heat reservoir to allow a specification of constant temperature instead. We now want to go one step further, bringing thermodynamics into the realm of chemical physics where most experiments are conducted at constant temperature and constant pressure. For this, we will see that the Gibbs free energy is the ideal thermodynamic potential.

5.1 Chemical Potentials for Gases

The **chemical potential** μ represents the energy required to introduce a single particle into the system, with no heat exchange or work done. Mathematically, this is written as:

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

where S is constant to prevent heat exchange and V is constant to prevent work. Since it is rather difficult to keep both of these extensive quantities constant in a real experiment, it is favourable to instead doubly Legendre transform the energy into the Gibbs free energy to obtain:

$$\mu = \mu(T, p, N) = \left(\frac{\partial G}{\partial N} \right)_{T, p} \quad (5.1)$$

In a real experiment, the temperature and pressure of the system may change over the course of an experimental run, but it is really just the initial and final values that matter since the Gibbs free energy is a state variable. The above definition now hints to us that the chemical potential in the natural variables (T, p, N) somehow represents the Gibbs free energy per particle in the system, and we can formalise this mathematically. We start by writing the defining extensive relation for the Gibbs free energy:

$$G(T, p, \lambda N) = \lambda G(T, p, N)$$

Differentiating this with respect to λ , we get:

$$\frac{d}{d\lambda} G(T, p, \lambda N) = G(T, p, N)$$

We then apply the chain rule to the left, and finally set $\lambda = 1$ to obtain:

$$\left(\frac{\partial G}{\partial N} \right)_{T, p} N = G(T, p, N) \implies \mu(T, p) = \frac{1}{N} G(T, p, N)$$

where the natural variables of μ are now only the temperature T and pressure p . This is a result of the extensivity of $G(T, p, N)$, combined with the fact that μ in these variables is obtained by simply dividing the

Gibbs free energy by N , implying that it cannot possibly contain any N dependence, otherwise it would no longer be intensive. We thus have a formal definition of the chemical potential in the Gibbs potential, where it is interpreted as the Gibbs free energy per particle at constant temperature and pressure. We can check this value for the ideal gas using its equation of state. We start with the fundamental relation in the Gibbs representation from Eq. 3.24, and divide through by N then set $dN = 0$ to write the Gibbs free energy per particle g , in terms of the entropy per particle s and volume per particle v (using lowercase variables):

$$dg = -s dT + v dp \implies dg = -s dT + \frac{k_B T}{p} dp$$

where we use the ideal gas equation of state. Integrating with respect to both variables gives:

$$g(T, p) = k_B T \log\left(\frac{p}{p_0}\right) + f(T)$$

where $f(T)$ is some unknown function of the temperature. We now use the fact that $g(T, p) \equiv \mu(T, p)$ to conclude that:

$$\mu(T, p) = k_B T \log\left(\frac{p}{p_0}\right) + f(T) \quad (5.2)$$

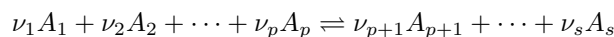
In this expression, one can only determine the exact expression if $f(T)$ is known through integration of the entropy per particle $s(T, p)$ ²⁵.

5.2 Chemical Reactions

In moving to our study of chemical physics, we must start considering systems with multiple species of particles. Suppose we have s species present in our system, then the fundamental relation in the Gibbs representation is:

$$G = G(T, p, \{N_j\}) \implies dG = -S dT + V dp + \sum_{j=1}^s \mu_j dN_j \quad (5.3)$$

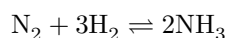
with the natural variables as (T, p, N_1, \dots, N_s) , which suggests that we should write the chemical potentials as $\mu_j(T, p, N_1, \dots, N_s)$, using two intensive variables and s extensive variables. We can instead convert this representation into entirely intensive parameters, which will then require one less dependency (as we established earlier with the Gibbs–Duhem relation). We define the **concentration** with respect to the s -th species as $c_j \equiv N_j/N_s$ as a set of $(s - 1)$ intensive parameters, where there is one less concentration than there are species since the last concentration is trivially unity. Thus, we obtain the chemical potential as a function of $(s + 1)$ intensive parameters: $\mu(T, p, c_1, \dots, c_{s-1})$. Now, in a general chemical reaction, we have some process:



where we have p reactants (on the left), q products (on the right) and the relation $p + q = s$. The species involved in the reaction are labelled by the set $\{A_j\}$, while their coefficients $\{\nu_j\}$ specify the relative number

²⁵Once we incorporate this with statistical mechanics, this is easily seen to be possible since the Sackur–Tetrode equation gives us the entropy of an ideal gas in terms of the natural extensive variables. Using the equations of state for T and p , this can be easily converted to give $s(T, p)$ which is then integrated to obtain the functional form of $f(T)$. Realistically, one would have a far easier time just looking up these functions in a table of standard Gibbs free energies.

of particles per reaction, also known as the **stoichiometric coefficients** of the reaction. As an example, we consider the following reaction for the production of ammonia, commonly known as the Haber–Bosch process²⁶:



In this reaction, we have $\{A_j\} = \{\text{N}_2, \text{H}_2, \text{NH}_3\}$, $\{\nu_j\} = \{-1, -3, 2\}$, $p = 2$, $q = 1$ and $s = 3$. Notice that $\nu < 0$ for reactants while $\nu > 0$ for products, indicating that reactants vanish while products appear in the course of the forward reaction. Thus, we can denote any general chemical reaction simply by:

$$\sum_{j=1}^s \nu_j A_j = 0 \quad (5.4)$$

In chemistry, equilibrium is typically considered to be measured at constant temperature and pressure, since reactions are most commonly carried out in an open environment. We thus start with the total differential for the Gibbs free energy (assuming constant T and p):

$$dG = \sum_{j=1}^s \mu_j dN_j = 0$$

where we set the differential to zero since the Gibbs free energy is minimised at equilibrium. Now, from Eq. 5.4 we note that the change in the particle number of the j -th species is directly proportional to ν_j , and this proportionality factor is simply the number of chemical reactions which took place in the system. We then denote the **number of chemical reactions** by \tilde{N} to write:

$$\begin{aligned} dG &= \sum_{j=1}^s \mu_j dN_j \\ &= d\tilde{N} \sum_{j=1}^s \mu_j \nu_j = 0 \end{aligned}$$

We can finally set the sum of this product to zero since $d\tilde{N}$ is arbitrary, giving us the condition for chemical equilibrium.

Definition 5.1: Condition for Chemical Equilibrium

Given a chemical reaction $\sum_{j=1}^s \nu_j A_j = 0$ where $\{A_j\}$ are the various particle species present (either reactants or products), the chemical potentials at equilibrium must satisfy:

$$\sum_{j=1}^s \mu_j \nu_j = 0 \quad (5.5)$$

where μ_j is the chemical potential of the j -th species and ν_j is its corresponding stoichiometric coefficient. The stoichiometric coefficients are positive for products and negative for reactants.

²⁶German chemists Fritz Haber and Carl Bosch received the Nobel Prize in Chemistry in 1918 for this work. It is arguably one of the most important industrialised chemical processes used today, responsible for the production of ammonia used in fertilisers and industrial feedstock for most of the world. While this work is highly respectable, I personally believe it should also be noted that immediately after development of this process, Haber spent the next several years leading up to and at the start of World War I developing methods for extracting large quantities of chlorine gas for use in trench warfare. He recruited large numbers of chemists and physicists, including the future Nobel laureates Otto Hahn, James Franck and Gustav Hertz, and led a team dedicated to the development of chemical weaponry. His work was posthumously used to also develop Zyklon B, the notorious chemical weapon of choice in the gas chambers during the Holocaust. For these reasons, Fritz Haber is dubbed the “father of chemical warfare”.

This is the general requirement for chemical equilibrium in a multi-species reaction, and provides a constraint amongst all the intensive chemical potentials at equilibrium²⁷.

5.2.1 Reactions in an Ideal Gas

We can study the idea of chemical equilibrium in a reversible reaction more specifically under the assumption that the reactants and products are ideal gases. While this is never true in reality, the result obtained here is an excellent approximation to the true observed behaviour since equations of state for actual reactants and products are typically of the form $f(p, V, T) = 0$. We first define the **partial pressure** p_j as the pressure produced by the j -th species if all other species' particles were removed, and the j -th species occupied the full volume of the container. Mathematically, we define it as:

$$p_j = \frac{N_j}{N}p \implies p_j V = N_j k_B T \quad (5.6)$$

We now take the chemical potential of an ideal gas from Eq. 5.2 and invoke the equilibrium condition from Eq. 5.5 to write:

$$\begin{aligned} \sum_{j=1}^s \nu_j k_B T \log p_j &= - \sum_{j=1}^s \nu_j f_j(T) \\ k_B T \log \left(\prod_{j=1}^s p_j^{\nu_j} \right) &= - \sum_{j=1}^s \nu_j f_j(T) \end{aligned}$$

where in the first line, the reference pressure p_0 is no longer necessary since this role is fulfilled by the total pressure p .

Definition 5.2: Law of Mass Action

At chemical equilibrium, the reactants and products in a system satisfy the **law of mass action**:

$$\prod_{j=1}^s p_j^{\nu_j} = \exp \left[- \frac{1}{k_B T} \sum_{j=1}^s \nu_j f_j(T) \right] \equiv K_p(T) = \frac{k_f}{k_b} \quad (5.7)$$

This determines the partial pressure of each species at equilibrium. The quantity $K_p(T)$ is commonly known in chemistry as the **equilibrium constant** of a reaction, and is typically something that can be looked up in a table. The final ratio k_f/k_b expresses the equilibrium constant as a ratio of **rate constants**, where k_f is known as the forward reaction rate constant, while k_b is the backward reaction rate constant. These rate constants determine the speed at which the forward and backward reactions occur, and they proceed in a manner at equilibrium which leaves the partial pressures unchanged.

More commonly the equilibrium constant is specified in terms of the **standard Gibbs free energy of reaction** ΔG^\ominus , defined as:

$$\Delta G^\ominus = -k_B T \log K_p(T) \stackrel{!}{=} \sum_{j=1}^s \nu_j f_j(T) \quad (5.8)$$

²⁷This condition is not satisfied when the system is out of equilibrium, since an imbalance in chemical potentials is required for particles to leave one state and enter another.

where the second equality is only true for an ideal gas reaction. The superscript Θ here refers to the fact that this is defined at some standard temperature and pressure, typically taken to be 298.15 K and 100 kPa (or 1 atm = 101.325 kPa in IUPAC-standardised references before 1982, or NIST references today). We can similarly formulate the law of mass action in terms of species concentration (for aqueous reactions) by defining a species concentration $c_j \equiv N_j/V = p_j/k_B T$. This gives:

$$\prod_{j=1}^s c_j^{\nu_j} = (k_B T)^{-\sum_j \nu_j} \exp \left[-\frac{1}{k_B T} \sum_{j=1}^s \nu_j f_j(T) \right] \equiv K_c(T) \quad (5.9)$$

where the equilibrium constant must be modified. To demonstrate how this enables a reaction to occur, we consider a product of concentrations such that $\prod_{j=1}^s p_j^{\nu_j} < K_p(T)$, so the system is not in equilibrium. This implies that an excess of the reactants exists (since the denominator is larger than its equilibrium value), and a deficiency of the products. The chemical potentials of the products then become negative, while those of the reactants remain positive (as seen from Eq. 5.2 with $p_0 = p_{\text{eq}}$), favouring the process in which particles are removed from the reactant state and added to the product state (i.e. the forward reaction). This continues until equilibrium is re-established, at which point we achieve $\prod_{j=1}^s p_j^{\nu_j} = K_p(T)$. This is an example of what you may have seen in high school as **Le Chatelier's principle**, whereby the system acts in a way such that the suitable reaction is favoured to re-establish equilibrium! In our thermodynamic picture, this process of re-equilibration following a perturbation is no magic, it is simply the act of the system minimising its Gibbs free energy!

5.3 The Heat of Reaction

We have just established a physical mechanism through which chemical reactions occur, and how they eventually reach chemical equilibrium while remaining robust to perturbations in the constituent particle numbers. The equilibrium constant $K_p(T)$ was also written as a function of the temperature, and this suggests that the chemical equilibrium state itself is temperature-dependent. To study this, we first consider the derivative:

$$\begin{aligned} \frac{\partial}{\partial T} \left(\frac{G}{T} \right)_{p,N} &= \frac{1}{T^2} \left[T \left(\frac{\partial G}{\partial T} \right)_{p,N} - G \right] \\ &= \frac{1}{T^2} (-TS - G) \\ &= -\frac{H}{T^2} \end{aligned}$$

where H is the enthalpy of the system. This gives us the following relation:

Definition 5.3: Heat of Reaction and the Gibbs–Helmholtz Equation

The enthalpy of a reaction is related to the change in Gibbs free energy of the system undergoing the reaction by the **Gibbs–Helmholtz equation**:

$$\Delta H = -T^2 \frac{\partial}{\partial T} \left(\frac{\Delta G}{T} \right)_{p,N} \equiv Q_p \quad (5.10)$$

This change in enthalpy is also commonly known as the **heat of reaction** Q_p , and is equal to the amount of heat released or absorbed in breaking and forming the chemical bonds as a result of the underlying forward and backward reactions.

We now return to the change in Gibbs free energy for a system undergoing a chemical reaction, and set $d\tilde{N} = 1$ to write:

$$\begin{aligned}\Delta G &= \sum_{j=1} \mu_j \nu_j \\ &= k_B T \sum_{j=1} \nu_j \log p_j - k_B T \log K_p(T)\end{aligned}$$

Finally, we insert this expression into the Gibbs–Helmholtz equation, noticing that the derivative kills the first term, leaving us with:

Definition 5.4: The Van t' Hoff Equation

Van t' Hoff's equation relates the change in the equilibrium constant $K_p(T)$ of a reaction to a change in the temperature T of the system, given the enthalpy of reaction ΔH .

$$\Delta H = k_B T^2 \frac{d}{dT} \log K_p(T) \quad (5.11)$$

Van t' Hoff's equation is powerful in formulating a physical picture of chemical reactions, in that it tells us which way a reaction proceeds when the temperature of the system is changed, since the reaction always proceeds in a manner to counteract the addition or removal of heat to a system in order to minimise its Gibbs free energy (and enthalpy). As an explicit demonstration, we consider a system with positive heat of reaction ($Q_p > 0$), so we have an endothermic reaction. Increasing the temperature then requires that $K_p(T)$ increase as well, since $\log K_p(T)$ is a monotonically increasing function of T . Then, from the expression for the equilibrium constant we see $K_p(T) = \prod_j p_j^{\nu_j}$ demands that the forward reaction should proceed to absorb this heat. This is yet another demonstration of Le Chatelier's principle, now for the case of a change in temperature!

6 Phases

In the previous sections, we had studied systems of increasing complexity, most recently relaxing the restriction of particle conservation. We now proceed one step further by introducing systems with several co-existing phases, alongside the possibility of more than one particle species. We define a **phase** to be some homogeneous part of a system that is separated via a boundary from other parts of the system (other phases). **Phase transitions** occur when the system changes from one phase to another at some suitable values of the system's parameters. Furthermore, we can treat each phase as an open subsystem of the whole, exchanging particles with its neighbouring subsystems in state space. First, we should get some mathematical notation introduced to deal with the phases. We denote the particle numbers of a multi-species multi-phase system by $\{N_j^{(k)}\}$, where j denotes the species and k denotes the phase. The fundamental relation in energy representation is then written as:

$$dE = T dS + \sum_{i=1}^r F_i dx_i + \sum_{j=1}^s \sum_{k=1}^t \mu_j^{(k)} dN_j^{(k)} \quad (6.1)$$

We now have $(1 + r + st)$ independent extensive variables, paired with corresponding intensive variables (we will re-formulate the dependency of intensive variables for multi-phase systems later). Performing a Legendre transformation, we end up with the Gibbs free energy as usual:

$$dG = -S dT - \sum_{i=1}^r x_i dF_i + \sum_{j=1}^s \sum_{k=1}^t \mu_j^{(k)} dN_j^{(k)} \quad (6.2)$$

which is still equally valid in the multi-phase context as long as we are keeping p and T constant (in the special case of $r = 1$). For this section, we will assume that each species is chemically inactive, with no ability to transmute amongst each other. We can then treat each phase within a species as an independent species, allowing us to write the Gibbs free energy of the total system (using the Euler equation) as:

$$G(T, \{F_i\}, \{N_j\}) = \sum_{j=1}^s \sum_{k=1}^t \mu_j^{(k)} N_j^{(k)}$$

Note that the chemical potential here is now a function of the natural variables:

$$\mu_j^{(k)} = \mu_j^{(k)}(T, \{F_i\}, N_1^{(k)}, \dots, N_{s-1}^{(k)}, N_s^{(k)})$$

Now, we use the same trick as before to turn the st extensive particle numbers into $(st - 1)$ intensive concentrations by dividing each particle number by $N_j^{(k)}$, giving us the chemical potentials as:

$$\mu_j^{(k)} = \mu_j^{(k)}(T, \{F_i\}, c_1^{(k)}, \dots, c_{s-1}^{(k)}) \quad (6.3)$$

6.1 Equilibrium Conditions for Multi-Phase Systems

The conditions of thermodynamic equilibrium set out in Def. 0.1 dictate how the intensive variables between two systems in contact must be equal. We later used this idea to develop the principle of Gibbs free energy minimisation. Now, with different phases in play, we need to re-develop the argument once more. At constant temperature and pressure, the first differential of the Gibbs free energy is:

$$dG = \sum_{j=1}^s \sum_{k=1}^t \mu_j^{(k)} dN_j^{(k)} = 0$$

Maintaining the assumption that the system's species are chemically inactive, we can denote the particles in the j -th species undergoing a phase transition $k \rightarrow l$ as $dN_j^{(k \rightarrow l)}$. Then, the total change in the k -th phase is:

$$dN_j^{(k)} = \sum_{l=1}^t dN_j^{(k \rightarrow l)} = - \sum_{l=1}^t dN_j^{(l \rightarrow k)}$$

with the obvious imposition that $dN_j^{(k \rightarrow k)} = 0$. At equilibrium, we then obtain:

$$\sum_{j=1}^s \sum_{k=1}^t \sum_{l=1}^t \mu_j^{(k)} dN_j^{(k \rightarrow l)} = 0$$

where we sum over all s species, all t initial phases and t final phases. Since this sum vanishes, we can simply add a copy of it with no consequence:

$$\sum_{j=1}^s \sum_{k=1}^t \sum_{l=1}^t \mu_j^{(k)} dN_j^{(k \rightarrow l)} + \sum_{j=1}^s \sum_{k=1}^t \sum_{l=1}^t \mu_j^{(l)} dN_j^{(l \rightarrow k)} = 0$$

where I have renamed the dummy variables $k \leftrightarrow l$ in the second term. Finally, we flip the direction of phase transfer in the second term to arrive at:

$$\sum_{j=1}^s \sum_{k=1}^t \sum_{l=1}^t (\mu_j^{(k)} - \mu_j^{(l)}) dN_j^{(k \rightarrow l)} = 0$$

Since this is true for any arbitrary set of $\{dN_j^{(k \rightarrow l)}\}$, the factor in parentheses must vanish. This leads us to the conclusion that $\mu_j^{(k)} = \mu_j^{(l)}$, for all pairs of phases (k, l) within each species j . This confirms our intuition that the chemical potential across all phases is equal at equilibrium. In a similar fashion, it can be shown that the temperatures and other intensive quantities must also be equal.

6.2 The Thermodynamic Stability Criterion

With extremal principles in place for the derivative of the thermodynamic potentials with respect to the extensive variables, we know what to expect in terms of the behaviour of these potentials as the system is perturbed slightly from its equilibrium point. Let us now extend this discussion by studying the conditions that must be obeyed by a system in a **stable thermodynamic equilibrium**. Knowing that a thermodynamic system contains an incredibly large number of particles ($N \gtrsim 10^{23}$), it is unavoidable that fluctuations will occur. These fluctuations refer to small deviations in the actual values of the characterising macroscopic parameters that we have used in specifying an equilibrium state (such as E, S, V , etc.). Stability here then refers to the idea that a system subject to these perturbing fluctuations will return to the equilibrium state, and not allow for unbounded growth of these microscopic fluctuations, destroying the equilibrium. These unbounded fluctuations are often seen in the vicinity of a phase transition, where the system macroscopically exhibits a marked change in behaviour.

For a system partitioned into two subsystems that are in thermodynamic equilibrium with one another, we have already established that the extensive variables add while the intensive variables are the same for each subsystem. The principle of energy minimisation then tells us that a small fluctuation in the entropy $\delta S_A = -\delta S_B$ in the subsystems must necessarily increase the internal energy of the total system: $\delta E > 0$. We formalise this by expanding such fluctuations to second order:

$$\begin{aligned} (\delta E)_{S,V,N} &= E_A(S_A + \delta S_A, V_A, N_A) - E_A(S_A, V_A, N_A) \\ &\quad + E_B(S_B + \delta S_B, V_B, N_B) - E_B(S_B, V_B, N_B) \\ &= \left(\frac{\partial E_A}{\partial S_A} \right)_{V_A, N_A} \delta S_A + \left(\frac{\partial E_B}{\partial S_B} \right)_{V_B, N_B} \delta S_B \\ &\quad + \frac{1}{2} \left(\frac{\partial^2 E_A}{\partial S_A^2} \right)_{V_A, N_A} (\delta S_A)^2 + \frac{1}{2} \left(\frac{\partial^2 E_B}{\partial S_B^2} \right)_{V_B, N_B} (\delta S_B)^2 + \mathcal{O}(\delta S^3) \\ &> 0 \end{aligned}$$

The energy minimisation principle ensures that the two first derivatives cancel each other out, so we must have the second derivatives obeying:

$$\begin{aligned} 0 &< \frac{1}{2} \left(\frac{\partial^2 E_A}{\partial S_A^2} \right)_{V_A, N_A} (\delta S_A)^2 + \frac{1}{2} \left(\frac{\partial^2 E_B}{\partial S_B^2} \right)_{V_B, N_B} (\delta S_B)^2 \\ &= \frac{1}{2} \left(\frac{\partial T_A}{\partial S_A} \right)_{V_A, N_A} (\delta S_A)^2 + \frac{1}{2} \left(\frac{\partial T_B}{\partial S_B} \right)_{V_B, N_B} (\delta S_B)^2 \end{aligned}$$

Since the partitioning is arbitrary, it follows that:

$$\left(\frac{\partial T}{\partial S}\right)_{V,N} > 0 \quad (6.4)$$

This is a necessary condition for stable thermal equilibrium. If this is not satisfied, certain small fluctuations of the entropy would decrease the internal energy of the system, spontaneously bringing it into a state with lower energy. This identifies the previous state of the system as an unstable equilibrium. From the definition of the heat capacity in Eq. 2.18, it is then easy to see that:

$$C_V = T \left(\frac{\partial S}{\partial T}\right)_{V,N} > 0 \quad (6.5)$$

Thus, the constant volume heat capacity of a system is strictly positive in a stable system at thermal equilibrium. Negative heat capacities would allow heat to spontaneously flow from the colder part of a system to a hotter part, increasing the internal temperature differential and moving it further from equilibrium. Clearly, the requirement for thermodynamic stability has profound implications on the response functions of a system, so we should explicitly look at how the others would also behave. Considering now fluctuations in the volume of a subsystem and applying the same reasoning as above leads us to the conclusion that:

$$\left(\frac{\partial^2 E}{\partial V^2}\right)_{S,N} = -\left(\frac{\partial p}{\partial V}\right)_{S,N} > 0 \quad (6.6)$$

From the definition of the compressibility in Eq. 1.7, we see that the adiabatic compressibility κ_S must obey:

$$\kappa_S = -\frac{1}{V} \left(\frac{\partial V}{\partial p}\right)_{S,N} > 0 \quad (6.7)$$

Intuitively, this makes sense since placing additional pressure on a system should decrease its volume instead of increasing it. If the opposite were true, the system would expand without bound until its volume eventually becomes comparable to that of the reservoir, possibly leading to an explosion. The adiabatic compressibility is only useful in situations where volume changes can be performed quickly, so we should instead aim to make a similar statement about the isothermal compressibility κ_T , where the system is held at a constant temperature. This requires us to instead work with the Helmholtz free energy since its natural variables are (T, V, N) . Fluctuations in the volume of a system then lead to second-order changes in the Helmholtz free energy given by:

$$0 < (\delta F)_{T,V,N} = \frac{1}{2} \left(\frac{\partial^2 F_A}{\partial V_A^2}\right)_{T_A, N_A} (\delta V_A)^2 + \frac{1}{2} \left(\frac{\partial^2 F_B}{\partial V_B^2}\right)_{T_B, N_B} (\delta V_B)^2$$

This gives us the requirement that:

$$\left(\frac{\partial^2 F}{\partial V^2}\right)_{T,N} = -\left(\frac{\partial p}{\partial V}\right)_{T,N} > 0 \implies \kappa_T > 0 \quad (6.8)$$

as expected. Interestingly, one can also show that this leads to the requirement that the constant pressure heat capacity $C_p > 0$ (try it yourself; use the definition of the heat capacities and the triple product rule). Knowing that the heat capacity difference obeys $C_p - C_V > 0$, we are also able to conclude that $\kappa_T > \kappa_S$. That is, it is easier to compress a system coupled to a heat reservoir than one that is thermally isolated.

6.3 Gibbs' Phase Rule

Let's now return to the simpler situation where $r = 1$ and consider a gaseous system with mechanical work only in the form of $(-p dV)$. Such a system will have $(2 + st)$ intensive variables (s species and t phases each), not all of which are independent. Recalling that the chemical potentials can be written in terms of $(s - 1)$ intensive concentrations (Eq. 6.3) for each phase (when chemical inactivity is assumed), we know that we have introduced one constraint from each phase k . Since there are t such phases, we have t constraints in the form of concentrations which trivially give unity.

We also know that within each species, the chemical potentials for the phases must be equal at equilibrium. Since we have $(t - 1)$ such independent chemical potentials remaining for each species, we have a total of $s(t - 1)$ new constraints. Adding up the total number of constraints, we have D remaining independent intensive variables given by²⁸:

Definition 6.1: Gibbs' Phase Rule

Gibbs phase rule dictates the number of independent intensive parameters D in a multi-species multi-phase system, given by:

$$D = 2 + st - t - s(t - 1) = 2 + s - t \quad (6.9)$$

for s chemically-inactive species spread amongst t phases per species. Demanding that $D \geq 0$ (since the number of degrees of freedom cannot be negative) tells us that for an s -species system, there can maximally only be t simultaneously co-existing phases:

$$p \leq 2 + s \quad (6.10)$$

As a simple example, we consider a single-species system ($s = 1$) and observe that there are only a maximum of three co-existing phases allowed at any given point in the intensive parameter space. This gives us the inequality:

$$D = 3 - p \geq 0$$

If only one phase ($p = 1$) exists in some region, we have $D = 2$ which tells us that there are two degrees of freedom (p, T) still available. This constitutes some general region in the intensive 2D state space that we can denote as a single phase. If two phases ($p = 2$) coexist in some region, we have $D = 1$ which tells us that this coexistence region has only one degree of freedom, and thus it must be a line in the intensive state space! This line can be parametrised as $p = p(T)$, and is commonly known as a **coexistence curve**. Finally, if three phases simultaneously coexist ($p = 3$), then $D = 0$ implies that this must be a single point in state space, known as the **triple point** of the system! These are shown in Fig. 12 for greater clarity. A well-known case of a triple point is that of water, found at 273.16 K and 611.657 Pa where ice, water and steam are all simultaneously in coexistence. What makes a triple point so unique is that miniscule changes in p or T will quickly move the system to a region containing only one phase, so it is a great temperature standard (if one knows the pressure to great accuracy). For this reason, the SI base unit of Kelvin was defined as 1/273.16-th of the triple point of water up until 2019, when it was redefined to use k_B instead.

²⁸Take as much time as you need to convince yourself of these counting arguments. This is an important point, and it should not be covered hastily.

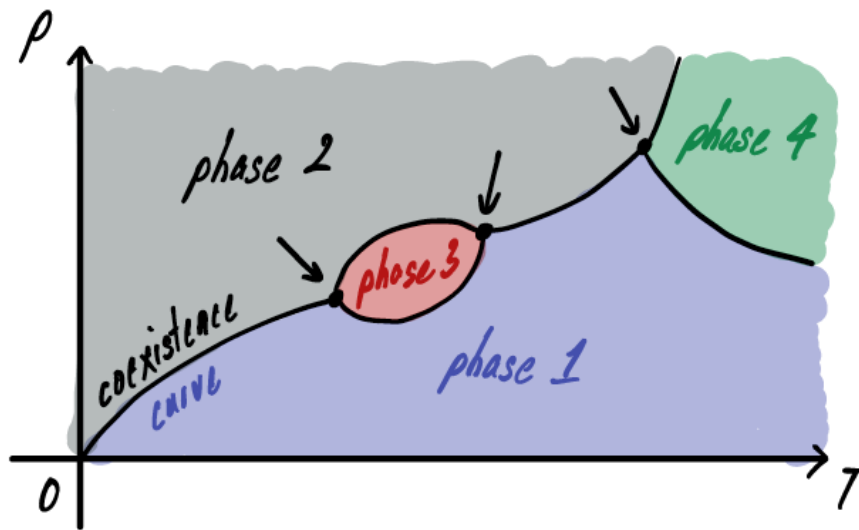


FIGURE 12: Gibbs' Phase Rule in the Intensive State Space. In a (p, T) state space for a single species ($s = 1$), the maximal number of coexisting phases allowed at any point is $p = 3$. This sets a constraint on the number of degrees of freedom D available to a region with p coexisting phases. Each individual phase ($D = 2, p = 1$) is a separately shaded region, while coexistence curves ($D = 1, p = 2$) are drawn as black lines and the triple points ($D = 0, p = 3$) are marked with black arrows.

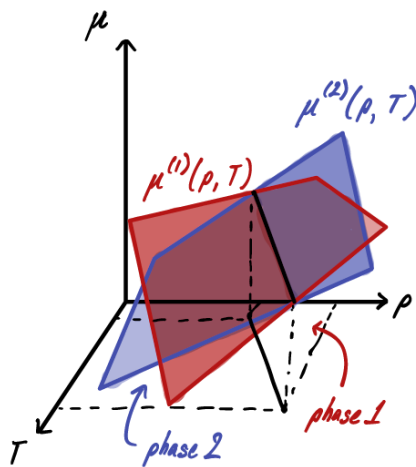


FIGURE 13: Gibbs Surfaces in the Intensive State Space. By drawing Gibbs surfaces corresponding to $\mu^{(k)}(p, T)$ above the (p, T) -plane, we can determine their curve of intersection to find the coexistence curve. Furthermore, the dominant phase observed in each region of the state space is given by the lower of the two surfaces.

Now, let's consider sitting on one of the coexistence curves where two phases are simultaneously coexisting in equilibrium. The (p, T) values along the line must be such that $\mu^{(1)}(p, T) = \mu^{(2)}(p, T)$, where $\mu^{(k)}(p, T)$ is the chemical potential as a function of temperature and pressure for the k -th phase in coexistence. If we now evaluate $\mu^{(1)}(p, T)$ for a set of (p, T) values in the region of phase 2 $\mathcal{P}^{(2)}$, we would obtain $\mu^{(1)}(p, T) > \mu^{(2)}(p, T)$. Similarly, in the region of phase 1 we would have $\mu^{(1)}(p, T) < \mu^{(2)}(p, T)$. This is because the chemical potential in some region is equal to the Gibbs free energy per particle, and since G is minimised at equilibrium, the phase with the smallest chemical potential is the actual phase observed in a given region. A geometric way to visualise the dominant phases is to draw surfaces of $\mu^{(k)}(p, T)$ at some height μ above the (p, T) -plane, known as **Gibbs surfaces**. The line of intersection between these planes represents the coexistence curve, while the dominant phase in each region is the phase with a lower value of $\mu^{(k)}(p, T)$. An illustration of the Gibbs surfaces is provided in Fig. 13.

It should also be noted that this extrapolation of $\mu^{(k)}(p, T)$ beyond its region of dominance (i.e. past the coexistence curve) is rather unphysical. In a real physical system, it is incredibly difficult to drive the system across the coexistence curve while

maintaining the previous phase since microscopic fluctuations in the particle number of one phase will unboundedly grow until the system macroscopically transitions to the new phase. One may be able to do this in

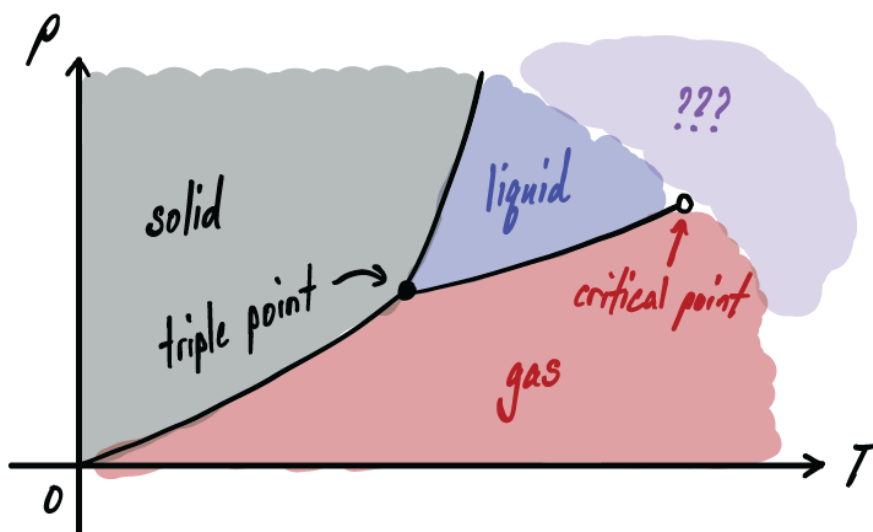


FIGURE 14: Critical Points in the Intensive State Space. In a (p, T) state space for a single species ($s = 1$), the coexistence curve bordering two phases can terminate at a critical point. In regions beyond the critical point, the distinction between the two previously coexisting phases is blurred, and one can traverse the state space from one phase to the other without ever undergoing a phase transition. In the above diagram, the liquid-gas coexistence curve terminates at a critical point, beyond which the purple region is indeterminate in terms of its phase.

small vicinities of the coexistence curve, however, such as in the familiar examples of superheated water (water existing above 100°C without turning to steam) or supercooled water (water existing below 0°C without turning to ice). These are simply metastable states, and will rapidly undergo a phase transition when subject to a modest perturbation.

The phase diagram of a realistic substance in typically accessible regimes of temperature and pressure²⁹ is not as complicated as the one shown in Fig. 6.9, and usually just involves three phases meeting at a triple point. However, real coexistence curves typically have an end beyond which they terminate. This point of termination is known as a **critical point**, and is illustrated in Fig. 14. Geometrically, coexistence curves lie on a cusp of the *minimal* Gibbs surface (the lower branch), but this cusp can eventually end. The upper and lower branches of the Gibbs surface beyond a critical point can no longer be distinguished from one another, having an appearance that looks similar to a multi-valued function in complex analysis. In this manner, one can traverse the state space going from one phase to another around a critical point, never having to undergo a phase transition by crossing the coexistence curve³⁰. We thus have to be careful about assigning a definite phase to regions of the intensive state space far from a coexistence curve.

6.4 The Clausius–Clapeyron Equation

To better understand the phase diagram, we should be able to exactly determine the equations of the coexistence curves that separate the various phases. We can do so by first considering the defining relation for a

²⁹Look up the real phase diagram of water and you’ll be surprised how incredibly complicated it is. There are six triple points and countless coexistence curves. Most of them are a result of ice having a highly nontrivial set of behaviours at higher pressures, with roughly 12 phases (that I am aware of).

³⁰Here’s an analogous situation — it is meaningful to say that Pennsylvania lies on the left of the Delaware River, while New Jersey lies on the right. It is not meaningful to say the same of Montreal or Paris, since the river terminates at some point in the continental USA. One could hop on a flight from Newark to Paris, then Paris to Montreal, and finally Montreal to Philadelphia, without having ever crossed the Delaware River (or cheating by going around the International Date Line).

coexistence curve (still assuming $s = 1$):

$$\mu^{(1)}(p, T) = \mu^{(2)}(p, T)$$

Now, we consider two infinitesimally separated points on the curve. These should satisfy:

$$\begin{cases} \mu^{(1)}(p, T) = \mu^{(2)}(p, T) \\ \mu^{(1)}(p + dp, T + dT) = \mu^{(2)}(p + dp, T + dT) \end{cases} \implies d\mu^{(1)} = d\mu^{(2)}$$

where we subtract the two equations to obtain the differential relation between the chemical potentials. Since the chemical potential is typically harder to work with, we want to convert this to the other intensive variables defining the state space. We use the Gibbs–Duhem relation to write:

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

Now, we apply the differential relation between the chemical potentials to obtain a differential relation between the temperatures and pressures as:

$$-\frac{S^{(1)}}{N^{(1)}} dT^{(1)} + \frac{V^{(1)}}{N^{(1)}} dp^{(1)} = -\frac{S^{(2)}}{N^{(2)}} dT^{(2)} + \frac{V^{(2)}}{N^{(2)}} dp^{(2)}$$

However, we know that along the curve the temperatures and pressures must change equally for both phases, so we can drop the superscripts on the differentials and group them. This leads us to write an equation for the derivative of the curve in (p, T) -space as:

$$\frac{dp}{dT} = \frac{-\frac{S^{(2)}}{N^{(2)}} - \frac{S^{(1)}}{N^{(1)}}}{\frac{V^{(2)}}{N^{(2)}} - \frac{V^{(1)}}{N^{(1)}}} = \frac{\Delta\left(\frac{S}{N}\right)}{\Delta\left(\frac{V}{N}\right)} = \frac{\Delta s}{\Delta v}$$

This finally leads us to the Clausius–Clapeyron equation:

Definition 6.2: Clausius–Clapeyron Equation

The coexistence curve between two phases in (p, T) -space is obtained by solving the differential equation:

$$\frac{dp}{dT} = \frac{\Delta s}{\Delta v} = \frac{L_{1 \rightarrow 2}}{T(V^{(2)} - V^{(1)})} \quad (6.11)$$

where $L_{1 \rightarrow 2}$ is known as the **latent heat of transformation** in going from phase 1 to phase 2, and is simply the heat involved in the entropy change when undergoing the phase transition. This is known as the **Clausius–Clapeyron equation**, and it details the slope of the coexistence curve for some transformation at any particular point along the curve in (p, T) -space.

Physically, the form of the Clausius–Clapeyron equation tells us why the melting-freezing coexistence curve is steeper than the condensation-vaporisation curve, since the volume change in the melting process is much smaller than that of the vaporisation process. The latent heat involved in breaking the bonds of a solid is also far greater than that of the bonds in a liquid phase, contributing to the steeper slope of the melting curve. In H_2O , this effect is exacerbated by the fact that the O–H bond lengths decrease upon melting, leading to a contraction of the fluid. This makes the slope of the melting curve negative, known as the **anomalous behaviour of water** at the melting transition.

We can also consider the vaporisation curve in going from the liquid to gas phase. This curve has a strictly positive slope, since the volume change in vaporisation is always positive. If we now assume that $V^{(2)} \gg V^{(1)}$, that the latent heat of transformation is constant in a small finite region, and that the vapour behaves as an ideal gas, we can use the ideal gas equation on the gaseous side of the phase transition to obtain an explicit equation for the curve. Invoking the first assumption ($V^{(2)} \gg V^{(1)}$) gives:

$$\frac{dp}{dT} = \frac{L_{1 \rightarrow 2}}{T(V^{(2)} - V^{(1)})} \approx \frac{L_{1 \rightarrow 2}}{TV^{(2)}}$$

Using the ideal gas equation of state then easily turns this into:

$$\frac{dp}{dT} = \frac{pL_{1 \rightarrow 2}}{k_B T^2}$$

This is integrated to give the liquid-gas coexistence curve as:

$$p(T) = p_0 \exp\left(-\frac{L_{1 \rightarrow 2}}{k_B T}\right) \quad (6.12)$$

which is typically valid for any gas phase that behaves like an ideal gas (negligible intermolecular interactions upon vaporisation). The latent heat of transformation can be easily looked up in a table of values.

6.5 Phase Transitions and the Maxwell Construction

In studying the various phases of a system and their coexistence regions thusfar, we have worked in a purely intensive state space of pressure p and temperature T . Previously, however, we have been used to working in the space of (p, V) , so it would be wise to revert back to that picture. For this, we first recognise that the volume V can be converted to an intensive variable $v = V/N$ known as the **specific volume**³¹. We then plot an isotherm in $p - v$ space, for some temperature T that cuts through the coexistence curve ($T < T_c$, where T_c is the temperature at the critical point) as shown in Fig. 15.

Here, $(p(T), T)$ is a point on the coexistence curve at some temperature T for which we draw the isotherm. The specific volumes $v^{(1)}$ and $v^{(2)}$ are the specific volumes of the two phases just at the edge of the coexistence region. Within each well-defined phase, the isotherm is continuous and has negative slope as dictated by the thermodynamic stability criterion in Eq. 6.7. We identify phase 1 with the gaseous phase and phase 2 with the liquid phase, since we expect the specific volume of the gaseous phase to be larger. Now, if we initially have some large specific volume $v > v^{(1)}(T)$ and reduce v isothermally (through contact with a large heat reservoir), the pressure will increase until we hit $v^{(1)}(T)$. Further reducing the volume at this point will not change the pressure, since all work done on the system goes directly into the latent heat of the transformation process. This continues until all the gas has condensed at which point we have the specific volume $v^{(2)}(T)$. Further volume reduction from here will return the standard behaviour of increasing the pressure.

While the Clausius–Clapeyron equation provides us with the parametrisation of the coexistence curve in (p, T) -space, it leaves us to deduce the specific volumes $v^{(1)}(T)$ and $v^{(2)}(T)$ ourselves. That is, we still need to determine the specific volumes at which the onset of coexistence occur when approaching the coexistence curve from either phase along an isotherm. Since the isotherm corresponds to a constant T (and constant N) process, we can write the pressure in (p, V) -space as a derivative of the Helmholtz free energy: $p = -(\partial F / \partial V)_{T, N}$. Notice that we can also remove the number dependence by switching to the intensive quantities:

³¹The specific volume is typically defined as $v = V/n$, where n is the number of moles of the gas present. We recognise that our definition in this review is simply related to this traditional definition by a factor of Avogadro's constant N_A .

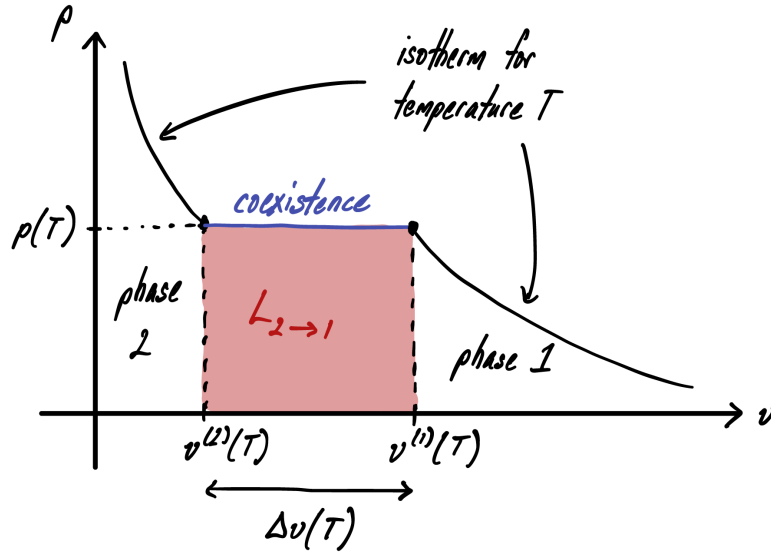


FIGURE 15: Coexistence Region Between Two Phases in $p - v$ Space. An isotherm occurring within a region passing through a coexistence curve in the $p - T$ space will exhibit a sudden jump in $p - v$ space. The discontinuity in the derivative of the isotherm represents the region where the temperature and pressure remain constant (on the coexistence curve) while all heat supplied to the system is used in enacting the phase transformation. This is represented by the shaded area, equal to the latent heat of transformation $L_{2 \rightarrow 1}$. In the region of coexistence, it is impossible to say for certain what phase the system is in, as it is almost always in a mixture of the two phases found on either side of the coexistence curve.

$$F(T, V, N) = NF\left(T, \frac{V}{N}, 1\right) = Nf(T, v)$$

where $f(T, v)$ is the Helmholtz free energy per particle. Through this, we can write the pressure as:

$$p = -\left(\frac{\partial F}{\partial V}\right)_{T, N} = -\left(\frac{\partial f}{\partial v}\right)_T = p(T, v)$$

Demanding that the pressures and chemical potentials (intersection of Gibbs surfaces) of the two phases on either side of the coexistence curve must be equal along the curve, we can write the set of equations:

$$\left(\frac{\partial f^{(1)}}{\partial v}\right)_T \Big|_{v=v^{(1)}} = \left(\frac{\partial f^{(2)}}{\partial v}\right)_T \Big|_{v=v^{(2)}} \quad (6.13)$$

$$\mu^{(1)}(T, v^{(1)}) = \mu^{(2)}(T, v^{(2)}) \quad (6.14)$$

where $f^{(k)}(T, v^{(k)})$ is the Helmholtz free energy per particle of the k -th phase at coexistence. The simultaneous solution of these equations for $v^{(1)}(T)$ and $v^{(2)}(T)$ gives us the specific volumes between which coexistence occurs for a given temperature T . At this stage, this sounds rather abstract, so let's try to unpack this more visually. Suppose we have the free energies for the two separate phases as functions $f^{(1)}(T, v^{(1)})$ and $f^{(2)}(T, v^{(2)})$. While in coexistence, Eq. 6.13 demands that the slope of $f^{(1)}$ at $v^{(1)}$ is equal to the slope of $f^{(2)}$ at $v^{(2)}$, as shown in Fig. 16.

Visually, we can see that the common tangent allows us to write:

$$f^{(2)}(T, v^{(2)}) - f^{(1)}(T, v^{(1)}) = -p(T)(v^{(2)} - v^{(1)})$$

Equivalently, we can group the terms for each phase to obtain:

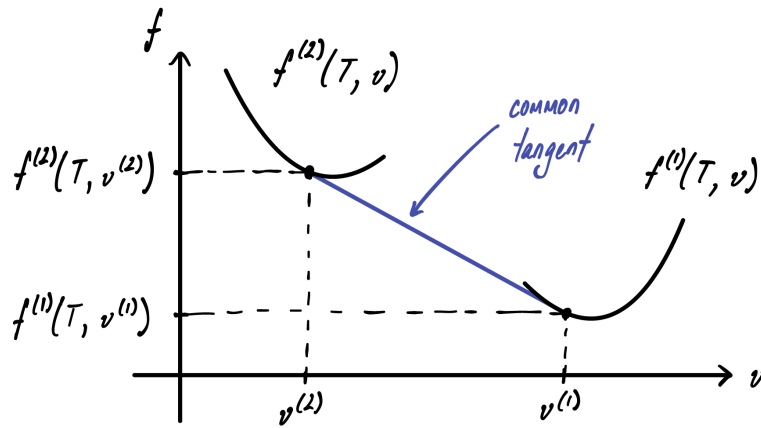


FIGURE 16: Common Tangent in Coexistence for Two Phases in $f - v$ Space. The coexistence region for two phases in the (f, v) space is determined by constructing a common tangent between the two free energy curves for the two phases. The equality of the slopes is imposed by Eq. 6.13.

$$(f + pv)^{(1)} = (f + pv)^{(2)}$$

Considering the Euler equation for the Gibbs free energy (or simply its definition as the Legendre transformation of the Helmholtz free energy on the (p, V) conjugate variable pair), we see that this condition is identical to Eq. 6.14, which sets the chemical potentials at coexistence to be the same! This confirms that the construction of the common tangent between the two phases' Helmholtz free energy curves is consistent with the condition of coexistence. Interestingly, the chemical potential is just the y -intercept of the common tangent, which can be easily seen considering the functional form of the tangent line (when viewed in the standard form of the expression for a straight line $y = mx + c$). The free energy in the coexistence region is then given by:

$$f(T, v) = f^{(1)}(T, v^{(1)}) \frac{v - v^{(2)}}{v^{(1)} - v^{(2)}} + f^{(2)}(T, v^{(2)}) \frac{v^{(1)} - v}{v^{(1)} - v^{(2)}} \quad (6.15)$$

where one should pay attention to the signs in the numerators. This is simply a convex sum of the two respective free energies on either side of coexistence, with the fractions given by the molar fraction of each phase. It is trivial to confirm that these molar fractions add to unity as one would expect.

When dealing with real systems, we will frequently run into situations where there is no known phase transition beforehand. Rather, we will have some expression for the equation of state or thermodynamic potential (such as the Helmholtz free energy) and we may want to make some predictions about the existence (or lack thereof) of a phase transition. In such situations, we may encounter a Helmholtz free energy that looks like that shown in Fig. 17.

In this diagram, we observe that the isotherm has the wrong sign in the range of v between $v^{(1)}(T)$ and $v^{(2)}(T)$. Specifically, we have:

$$\left(\frac{\partial^2 f}{\partial v^2} \right)_T < 0 \iff \left(\frac{\partial p}{\partial v} \right) > 0$$

which violates the thermodynamic stability criterion set out in Eq. 6.7. We rectify this by constructing the tangent between those two specific volumes, and this tangent now predicts the Helmholtz free energy in that range instead of the original equation of state. These are, once again, determined by solving the set of

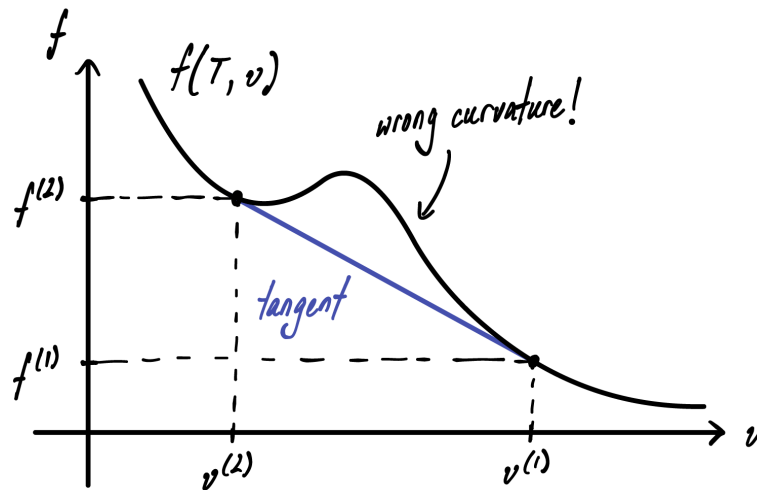


FIGURE 17: Instability in the Helmholtz Free Energy. In studying an unknown system, we may encounter an equation of state that produces a Helmholtz free energy which looks like this. In the region between $v^{(1)}(T)$ and $v^{(2)}(T)$, notice how the Helmholtz free energy is convex instead of concave. This suggests that the system as predicted by the equation of state is thermodynamically unstable in said region, and is likely to be undergoing a phase transition.

equations Eq. 6.13 and 6.14 for $v^{(1)}(T)$ and $v^{(2)}(T)$, with $p(T)$ providing the slope of the curve and μ providing the y -intercept.

Similarly, the thermodynamic instability can be seen on the $p - v$ diagram in Fig. 18 where it presents itself as a region in which the isothermal $p(T, v)$ curve is increasing instead of decreasing. We note that the construction of the common tangent in Fig. 17 demands that the free energy changes linearly with specific volume. This implies that the pressure must be a constant over the same region. Setting the pressure to a constant $p(T)$ in the region between $v^{(1)}(T)$ and $v^{(2)}(T)$ essentially truncates two curved shaded regions (sometimes called the *unphysical oscillations* of the equation of state). This procedure is known as a **Maxwell construction**³², where the pressure $p(T)$ is set such that the areas of the shaded regions above and below the line of constant pressure is equal. The physical justification for this is that the work done by compressing the gas through some specific volume should still be correct as predicted by the equation of state, since instability only predicts a change in phase but not a change in the internal energy of the substance. We thus require the following mathematical relation is obeyed:

$$\begin{aligned} (v^{(1)} - v^{(2)})p(T) &= f(T, v^{(2)}) - f(T, v^{(1)}) \\ &= \int_{v^{(2)}}^{v^{(1)}} dv \left(-\frac{\partial f}{\partial v} \right)_T \end{aligned} \quad (6.16)$$

The left hand side is simply the area under the rectangle defined by the boundaries $p(T)$, $v^{(2)} < v < v^{(1)}$ on the $p - v$ diagram. The right hand side is the area under the graph of the $p - v$ diagram in the same region of specific volume. These two terms are equal when the shaded areas are exactly equal. This condition for the Maxwell construction thus sets a sufficient number of constraints that enables us to determine the coexistence region from the equation of state alone.

³²It turns out that Gibbs also had a very similar construction, but he had yet to publicly speak of it in the same way that Maxwell did. He did, however, credit Gibbs with his construction when giving his first lecture on the subject of studying phase transitions in gaseous systems.

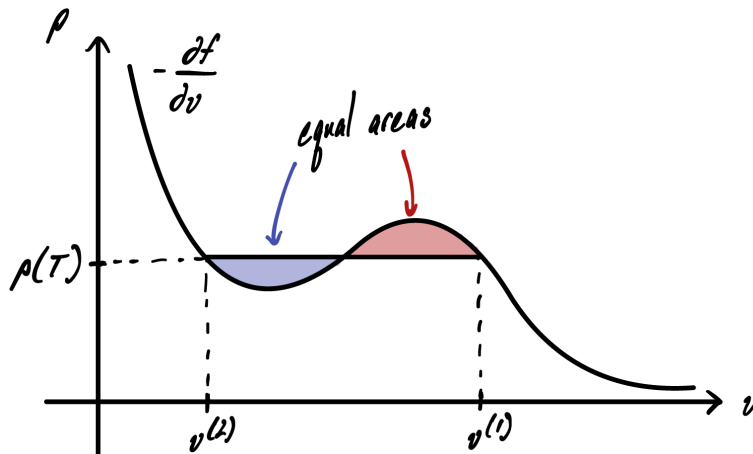


FIGURE 18: Maxwell Construction for an Isotherm with a Phase Transition. In studying an unknown system, we may encounter an equation of state that produces a Helmholtz free energy which looks like this. In the region between $v^{(1)}(T)$ and $v^{(2)}(T)$, notice how the Helmholtz free energy is convex instead of concave. This suggests that the system as predicted by the equation of state is thermodynamically unstable in said region, and is likely to be undergoing a phase transition.

6.6 The van der Waals Gas

As a final demonstration, let's apply what we have just learnt about phase transitions to the **van der Waals gas**. This is an approximate model for a real gas, with the equation of state:

$$\left(p + \frac{a}{v^2}\right)(v - b) = RT \quad \text{with } a, b > 0 \quad (6.17)$$

where we use the molar gas constant R instead of the Boltzmann constant k_B since this is how Johannes Diderik van der Waals first presented it³³. The equation is also written in purely intensive form, using the intensive specific volume v instead of the extensive volume V . This differs from our familiar ideal gas equation in two ways:

1. Accounting for the finite volume occupied by the constituent particles of a gas by ensuring that the minimal specific volume is $b > 0$.
2. Phenomenological introduction of the "kinetic pressure" av^{-2} to account for the resistance of molecules against externally applied pressure.

While the ideal gas equation of state predicts a gaseous state for the entire parameter space (since $\partial p/\partial v < 0$ everywhere) and hence no phase transitions, these two simple modifications add such great richness to the model of a real gas that it now exhibits a phase transition. This is easily shown by proving that there exists some region of the intensive state space in which the derivatives disobey the thermodynamic stability criterion. To demonstrate this, we first consider the derivatives of the equation of state in the more tractable form:

$$p(T, v) = \frac{RT}{v - b} - \frac{a}{v^2} \quad (6.18)$$

We know that a coexistence region exists for some range of v when we are within a suitable region of the (p, T) phase diagram. Specifically, for $T > T_c$ all the isotherms are stable over all v , while they start exhibiting

³³van der Waals received the Nobel Prize in Physics in 1910 for this work, amongst other developments in studying equations of state for fluid systems. This was also primarily the work which he conducted for his doctoral thesis entitled *O ver de continuïteit van den gas- en vloeistoftoestand Academisch proefschrift*. I miss the days when people won Nobels for their doctoral work.

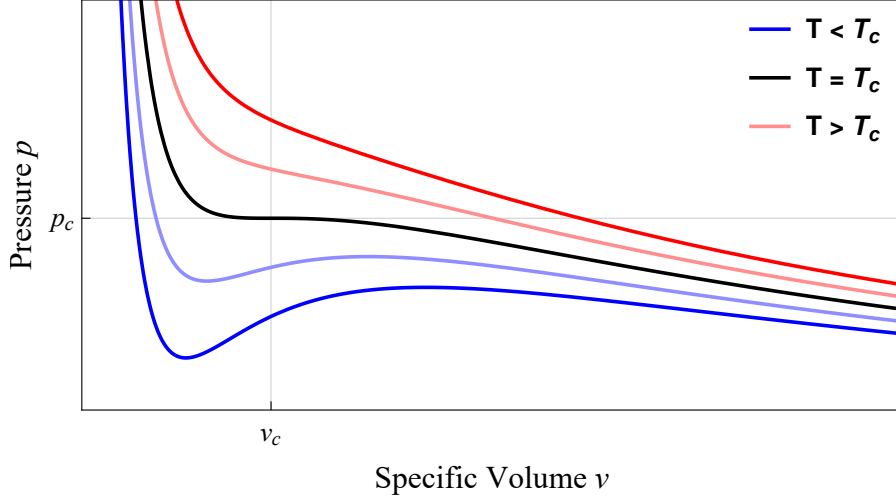


FIGURE 19: Critical and Near-Critical Isotherms of the van der Waals Gas. Above the critical temperature T_c , the isotherms are always stable since they obey the thermodynamic stability criterion. Below the critical temperature, they exhibit some region of instability over v . Thus, the critical isotherm for $T = T_c$ is found by demanding that both the first and second derivatives of $p(T, v)$ with respect to v vanish at the critical point.

an unstable region for $T < T_c$. To solve for this critical temperature T_c , and hence the critical pressure p_c and critical volume v_c , we demand that the isotherms first exhibit a stationary point of inflection once this critical temperature is crossed, as illustrated in Fig. 19. To show the existence of an inflection point, we find the derivatives:

$$\begin{aligned} \left(\frac{\partial p}{\partial v}\right)_T &= -\frac{RT}{(v-b)^2} + \frac{2a}{v^3} \\ \left(\frac{\partial^2 p}{\partial v^2}\right)_T &= \frac{2RT}{(v-b)^3} - \frac{6a}{v^4} \end{aligned}$$

Setting both of these derivatives to zero gives the condition for inflection, returning the critical parameters:

$$RT_c = \frac{8a}{27b}, \quad v_c = 3b, \quad p_c = p(T_c, v_c) = \frac{a}{27b^2} \quad (6.19)$$

Let's consider a subcritical isotherm $T < T_c$ and examine the Maxwell construction required for the range $v^{(2)} < v < v^{(1)}$. At the coexistence pressure $p(T)$, we have the equation:

$$p(T) = \frac{RT}{v^{(1)} - b} - \frac{a}{v^{(1)2}} = \frac{RT}{v^{(2)} - b} - \frac{a}{v^{(2)2}} \quad (6.20)$$

which forms two equations for the three unknowns that we are solving for ($v^{(1)}$, $v^{(2)}$ and $p(T)$). The third equation is provided by the defining condition for the Maxwell construction (Eq. 6.16), which gives:

$$\begin{aligned} (v^{(1)} - v^{(2)})p(T) &= \int_{v^{(2)}}^{v^{(1)}} dv \left(\frac{RT}{v-b} - \frac{a}{v^2} \right) \\ &= RT \log \left(\frac{v^{(1)} - b}{v^{(2)} - b} \right) - a \frac{v^{(1)} - v^{(2)}}{v^{(1)}v^{(2)}} \\ \implies p(T) + \frac{a}{v^{(1)}v^{(2)}} &= \frac{RT}{v^{(1)} - v^{(2)}} \log \left(\frac{v^{(1)} - b}{v^{(2)} - b} \right) \end{aligned} \quad (6.21)$$

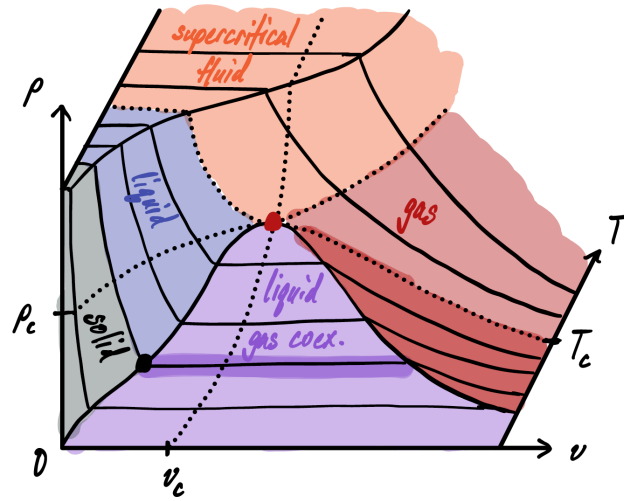


FIGURE 20: Phase Diagram of the van der Waals Gas. In the full intensive state space, we see that the onset of coexistence represents a sort of “opening up” of the isotherms to give way to an extended region of liquid-gas coexistence (purple). For subcritical temperature regions below the critical pressure (dark red), we say that the gas is in a saturated vapour state. In the region beyond the critical pressure, temperature and volume, the van der Waals fluid is said to be supercritical and there is no clear distinction between the liquid and gas phases.

Plugging Eq. 6.20 into Eq. 6.21, we turn the left hand side into:

$$\begin{aligned} p(T) + \frac{a}{v^{(1)}v^{(2)}} &= \frac{RT}{v^{(1)} - b} + \frac{a}{v^{(1)2}v^{(2)}} (v^{(1)} - v^{(2)}) \\ &= \frac{RT}{v^{(2)} - b} - \frac{a}{v^{(2)2}v^{(1)}} (v^{(1)} - v^{(2)}) \end{aligned}$$

From this, we can extract RT and $p(T)$ as:

$$RT = \frac{a}{[v^{(1)}v^{(2)}]^2} (v^{(1)} + v^{(2)}) (v^{(1)} - b) (v^{(2)} - b) \quad (6.22)$$

$$p(T) + \frac{a}{v^{(1)}v^{(2)}} = \frac{a}{[v^{(1)}v^{(2)}]^2} \left[(v^{(2)} - b)v^{(1)} + (v^{(1)} - b)v^{(2)} \right] \quad (6.23)$$

Combining all the intermediate equations thusfar gives us a self-consistent condition on the pairs of specific volumes between which coexistence occurs for a fixed temperature as:

$$\frac{v^{(1)}}{v^{(1)} - b} + \frac{v^{(2)}}{v^{(2)} - b} = \frac{v^{(1)} + v^{(2)}}{v^{(1)} - v^{(2)}} \log \left(\frac{v^{(1)} - b}{v^{(2)} - b} \right) \quad (6.24)$$

At this point, most textbooks would tell the reader that we have done all we can and that this should be solved numerically to obtain the coexistence volumes, from which the coexistence pressure can be determined for each Maxwell construction at a fixed temperature. It turns out that this can be solved analytically [9], but I am genuinely not willing to do that here since it does not aid the discussion³⁴.

The phase diagram that we have been looking at in $p - v$ space is merely a slice of the full (p, v, T) phase diagram shown in Fig. 20. The isotherms for the van der Waals gas shown in Fig. 19 represent what one would

³⁴Please try this for yourself if you are interested, it’s a good exercise in algebra (albeit a hard one since Maxwell himself was not able to think of how to do it).

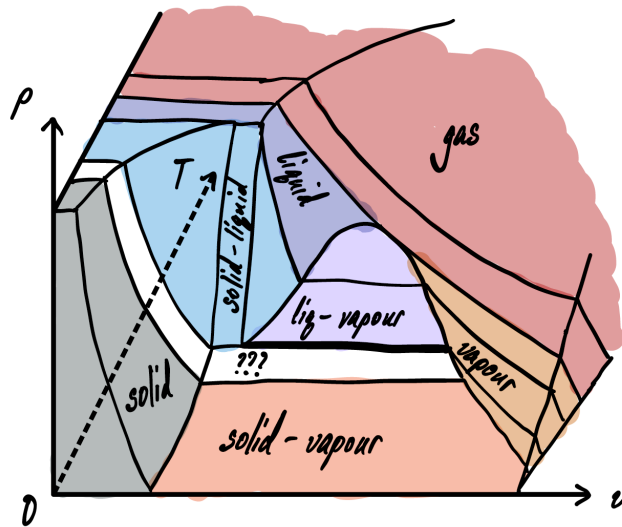


FIGURE 21: Schematic Phase Diagram of Water. The van der Waals equation of state generally captures the essence of the phase diagram of a substance like water, reproducing the general regions as observed experimentally. In this diagram, the supercritical region is labelled as a gas, but it should be noted that this is really a supercritical and unstable fluid. There is also a distinction made between gas and vapour here, where vapour is defined as a semi-gaseous state of liquid origin, whereas a gas is closer to the van der Waals model.

see if we looked from the front face of this 3D plot, where the subcritical isotherms have had their Maxwell constructions applied. The typical picture that we are used to from Fig. 14 is simply the view that one would have when looking at this plot from the right face.

It is rather incredible that an equation of state with just two phenomenological modifications (as in van der Waals' case) is a surprisingly accurate descriptor of most real gases. The general structure of the intensive state space is captured by this model, and it only fails in terms of numerics, where the critical parameters and coexistence curves have deviations due to the nature of the interactions between the specific molecules in a gas. There are other gas equations of state (Berthelot, Redlich-Kwong, Dieterici, etc.) which can be solved similarly to study the critical phenomena that they aim to model, but the van der Waals model is the simplest (and hence it is the one used in undergraduate thermal physics courses). These equations of state are typically semi-empirical in that they require some correction motivated by experiment, but provide a much more accurate numerical characterisation of the critical phenomena. Interactions are more accurately studied using cluster expansion models such as the virial equation of state, which accounts for the interactions between pairs of molecules, triples of molecules, etc. These are typically not solvable analytically, and are treated up to a truncated order in a perturbing parameter such as the interaction strength (or the density).

References

1. Nernst, W. The New Heat Theorem: its Foundations in Theory and Experiment. **Nature** **119**, 43–44 (Jan. 1927).
2. Clausius, R. **The Mechanical Theory of Heat** (Cornell University Library, 1879).
3. Carnot, N. L. S. **Reflections on the Motive Power of Fire (translated)** (ed Thurston, H.) (Dover Books, 1890).
4. Carathéodory, C. Untersuchungen über die Grundlagen der Thermodynamik. **Mathematische Annalen** **67**, 355–386 (Sept. 1909).
5. Callen, H. B. **Thermodynamics and an Introduction to Thermostatistics** 493. ISBN: 0471862568 (Wiley, 1985).
6. Lieb, E. H. & Yngvason, J. The physics and mathematics of the second law of thermodynamics. **Physics Reports** **310**, 1–96 (Mar. 1999).
7. Gibbs, J. W. **Elementary Principles in Statistical Mechanics** 224. ISBN: 9780486789958 (Dover Publications, Incorporated, 2014).
8. Sekerka, R. F. **Thermal Physics: Thermodynamics and Statistical Mechanics for Scientists and Engineers. Thermodynamics and Statistical Mechanics for Scientists and Engineers** 610. ISBN: 9780128033043 (Elsevier Science & Technology Books, 2015).
9. Englert, B.-G. **Lectures on Statistical Mechanics** (WORLD SCIENTIFIC, June 2020).