Statistical Physics (PHY831): Part 1 - The foundations

Phillip M. Duxbury, Fall 2011

Course Outline

Part 1: (LL, PB) Foundations: (11 lectures)
History of thermodynamics, kinetic theory and equilibrium statistical mechanics. Axioms of equilibrium statistical physics, Master equation, Liouville equation. Equilibrium conditions, Boltzmann weights and partition functions. Free energies and relation to partition functions. Survey of many particle, spin and field systems: the case of classical particles, the case of a non-interacting (Gaussian) polymer chain, the case of "classical" spins. Computational methods: molecular dynamics and Monte Carlo. Advanced thermodyamics, Maxwell relations and response functions. Fluctuations and relation to response functions.
Midterm 1, Lecture 11 (Monday Sept. 26)

Part 2: (H, PB, LL) Key solvable systems: (11 lectures)
Classical ideal gas, Non-interacting spin systems, Classical harmonic oscillators, Energy levels of a non-relativistic and relavistic particle in a box, ideal Bose and Fermi gases. Applications to atom traps, white dwarf and neutron stars, electrons in metals, photons and solar energy, phonons, Bose condensation and superfluidity, the early universe.
Midterm 2, Lecture 22 (Friday Oct. 21)

Part 3: (H, PB) Interacting systems, phase transitions and critical phenomena (9 lectures)
Interacting spin systems, Ising model. Interacting classical gas, cluster expansion, van der Waals equation of state, Virial Expansion. Superfluidity and superconductivity, phonons, rotons, vortices, BCS theory.
Midterm 3, Lecture 32 (Monday Nov. 14)

Part 4: (H, PB) Scaling and complex systems (9 lectures)
Phase transitions and critical phenomena, scaling, Landau theory, GL theory, RG. Equilibrium and non-equilibrium dynamics. Fractals, percolation, DLA, self-organized criticality, scale free networks
Midterm 4, Lecture 42 (Friday Dec. 9)

I. ORIGINS OF THERMODYNAMICS, KINETIC THEORY AND STATISTICAL MECHANICS

A. Engines

Thermodynamics originated with the study of heat engines. A precursor to heat engines, the vacuum pump, was designed by the German scientist Otto von Guericke who, in 1650, created the world's first ever vacuum in what are known as the Magdeburg hemispheres. He was driven to make a vacuum in order to disprove Aristotle's long-held supposition that "Nature abhors a vacuum".

Shortly thereafter, Irish physicist and chemist Robert Boyle learned of Guericke's designs and in 1656, in coordination with English scientist Robert Hooke, built an air pump. Using this pump, Boyle and Hooke noticed the pressure-volume correlation: $PV = constant$. At that time, air was assumed to be a system of motionless particles rather than a system of moving molecules. Later, after the invention of the thermometer, the temperature could be quantified, providing Gay-Lussac the opportunity to derive his law ($P \propto T$), which quickly led to the ideal gas law.

$$PV = Nk_BT = nRT \tag{1}$$

This law works quite well for most gases and even for steam, provided they are at pressures and temperatures well away from any phase transitions. The ideal gas law was thus the basis for a lot of the early heat engine development.

Even before the establishment of the ideal gas law, an associate of Boyle's named Denis Papin in 1679 built a high pressure steam cooker, also called a bone digester or pressure cooker, which is a closed vessel with a tightly fitting lid that confines steam until a high pressure is generated. Later designs included a steam release valve to keep the machine from exploding. By watching the steam release valve rhythmically move up and down at high heat levels, Papin conceived and designed a piston and cylinder engine. He did not however follow through with his design. Nevertheless, in 1697, based on Papins designs, engineer Thomas Savery built the first heat engine. Although these early heat or steam engines were crude and inefficient, they attracted the attention of the leading scientists of the time. One such scientist was Sadi Carnot (1796-1832), the father of thermodynamics, who in 1824 published "Reflections on the Motive Power of Fire", a discourse on heat, power, and engine efficiency. This marks the start of thermodynamics as a modern science.

Prior to 1697 and the invention of the Savery Engine, horses were used to power pulleys, attached to buckets, which lifted water out of flooded salt and coal mines in England. In the years to follow, more variations of steam engines were built, such as the Newcomen Engine, and later the Watt Engine. In time, these early engines would be utilized in place of horses. Thus, each engine began to be associated with a certain amount of "horse power" depending upon how many horses it had replaced. The main problem with these first engines was that they were slow and clumsy, converting less than 2% of the input fuel into useful work. Hence the need for a new science of engine dynamics was born. Carnot defined "motive power" to be the expression of the useful "effect" that a motor is capable of producing. He introduced the first modern day definition of "work" as weight lifted through a height. In 1769 Frenchman Nicolas-Joseph Cugnot build the first self-propelled steam engine. In 1804 Richard Trevithick built the first steam locomotive and in 1829 George Stevenson build the Rocket a practical steam locomotive. This ignited the railway revolution in Europe and the US. Steam engines use the Rankine cycle and are used in most of the electricity production in the world, including coal, gas, nuclear and solar thermal generators

In 1843, James Joule experimentally found the mechanical equivalent of heat, and in 1845 he reported his best-known experiment, involving the use of a falling weight to spin a paddle-wheel in an barrel of water, heating the water and allowing him to estimate a mechanical equivalent of heat. This led to the theory of conservation of energy and explained why heat can do work. The name "thermodynamics" was introduced, in 1849, by the British mathematician and physicist William Thomson (Lord Kelvin) who used the term in a paper on the "Carnot" efficiency of steam engines. These developments lead to the simple formula for the Carnot efficiency of an ideal heat engine,

$$\eta = \frac{W}{Q_h} = 1 - \frac{Q_c}{Q_h} = 1 - \frac{T_c}{T_h} \tag{2}$$

where W is the mechanical work done by the heat engine, Q_h is the heat energy supplied by an isothermal heat source with T_h it's temperature, while Q_c is the waste heat delivered to an isothermal cold reservoir that is at temperature T_c.

B. Introduction of relation between entropy and heat

In 1850, the famed mathematical physicist Rudolf Clausius defined the term entropy S to be the heat lost or turned into waste, stemming from the Greek word entrepein meaning to turn. The Clausius equality states that for a reversible process, we have,

$$\oint \frac{dQ}{T} = 0 \tag{3}$$

so that dQ/T is a state function, like internal energy, pressure and other thermodynamic functions. This lead him to introduce the idea that there is a new thermodynamic state function $S = Q/T$ that he called entropy. Note that the product $dQ = TdS$ has units of energy. Clausius also derived the virial theorem (1870) and the equipartition theorem follows from it. For an ideal gas the internal energy $U = \frac{1}{2}k_BT$ per degree of freedom (DOF).

C. Laws of thermodynamics

The laws of thermodynamics are based on a continuum "caloric theory" which did not explicitly require atomic scale understanding. In modern form the laws of thermodynamics, are:

Zeroth Law: If system A and system B are in thermal equilibrium with system C, then system A is in thermal equilibrium with system B.

First Law: Increase in internal energy of a system (ΔU) = heat supplied to the system (Q) - work done by the system W. This is often written in differential form $dU = dQ - dW$.

Second Law: When two isolated systems in separate but nearby regions of space, each in thermodynamic equilibrium in itself (but not necessarily in equilibrium with each other at first) are at some time allowed to interact, breaking the isolation that separates the two systems, allowing them to exchange matter or energy, they will eventually reach a mutual thermodynamic equilibrium. The sum of the entropies of the initial, isolated systems is less than or equal to the entropy of the final combination of exchanging systems. In the process of reaching a new thermodynamic equilibrium, total entropy has increased, or at least has not decreased. For reversible heat transfer $\delta Q = TdS$.

Third Law: The entropy of a perfect crystal at absolute zero is exactly equal to zero.

D. Key engine cycles and properties

Three ideal cycles that many engines aspire to are the Rankine cycle, the Carnot cycle and the Stirling cycle.

The Rankine cycle applies to ideal steam turbines: 1-2 water is pumped from low to high pressure; 2-3 an external heat source converts the water to (superheated) steam at constant pressure; 3-4 the steam expands driving a turbine. In an ideal cycle this is isentropic; 4-1, the steam is condensed to water (isothermal). Note that a reversible adiabatic process is isentropic.

The Carnot cycle applies to ideal IC engines: 1-2 heat supplied to produce isothermal expansion of gas at T_H; 2-3, isentropic expansion of gas, while cooling to T_C; 3-4 isothermal compression of gas at T_C with heat loss to reservoir; 4-1 isentropic compression of gas to temperature T_H.

Stirling engine applies to ideal heat pumps and refrigerators: 1-2, an external heat source produces isothermal expansion at T_H; 2-3, isovolumetric or isochoric heat removal; 3-4, isothermal compression with heat loss to a reservoir at temperature T_C; 4-1, isovolumetric or isochoric heat addition.

Carnot's theorem states:

- All irreversible heat engines between two heat reservoirs are less efficient than a Carnot engine operating between the same reservoirs.
- All reversible heat engines between two heat reservoirs are equally efficient with a Carnot engine operating between the same reservoirs.

The Carnot and Stirling engines have equal ideal efficiencies. The Rankine cycle is less efficient, but is more practical for many real applications, such as electricity generation.

E. History of kinetic theory and its relation to heat

Kinetic theory, the study of the motion of atomic scale particles, provides an atomic scale basis for entropy and heat and hence is an integral part of statistical mechanics. The idea that heat is a form of motion is perhaps an ancient one and is certainly discussed by Francis Bacon in 1620 in his Novum Organum. The first written scientific reflection on the microscopic nature of heat is probably to be found in a work by Mikhail Lomonosov (1711-1765), in which he wrote, "(..) movement should not be denied based on the fact it is not seen. Who would deny that the leaves of trees move when rustled by a wind, despite it being unobservable from large distances? Just as in this case motion remains hidden due to perspective, it remains hidden in warm bodies due to the extremely small sizes of the moving particles. In both cases, the viewing angle is so small that neither the object nor their movement can be seen."

During the same years, Daniel Bernoulli published his book "Hydrodynamics" (1738), in which he derived an equation for the pressure of a gas considering the collisions of its atoms with the walls of a container. He proved that this pressure is two thirds the average kinetic energy of the gas in a unit volume; and derived the relation between pressure, velocity and height in an incompressible fluid flow

$$P + \frac{1}{2}\rho v_D^2 + \rho g h = constant\ (incompressible\ fluid), \tag{4}$$

where P is pressure $\rho v_D^2/2$ is kinetic energy density of the moving fluid, ρ is fluid density, and h is height of the fluid in a gravitational field. Note that the kinetic energy in this expression does not include the random thermal motion of the atoms in the fluid. Instead v_D in this expression is more analogous to a drift velocity, and I have added the subscript to make this explicit, as students often confuse these two velocities. Bernoulli's ideas, however, made little impact on the dominant theory of heat (or caloric) engines. Bernoulli made a connection with Gottfried Leibniz's vis viva principle, an early formulation of the principle of conservation of energy, and the two theories became intimately entwined throughout their history. In addition to their conflict over the invention of calculus, Leibniz's vis viva theory seemed in opposition to the conservation of momentum introduced by Newton and Descartes, though now of course we know they are complementary. Though Benjamin Thompson suggested that heat was a form of motion as a result of his experiments in 1798, no attempt was made to reconcile the theoretical and experimental approaches, and it is unlikely that he was thinking of the vis viva principle.

John Herapath later independently formulated a kinetic theory in 1820, but mistakenly associated temperature with momentum rather than vis viva or kinetic energy. His work ultimately failed peer review and was neglected. John James Waterston in 1843 provided a largely accurate account, again independently, but his work received the same reception, failing peer review even from someone as well-disposed to the kinetic principle as Humphry Davy. Further progress in kinetic theory started only in the middle of the 19th century, with the works of Rudolf Clausius, James Clerk Maxwell, and Ludwig Boltzmann. In his 1857 work on the nature of the motion called heat, Clausius for the first time clearly states that heat is related to the average kinetic energy of molecules. This interested Maxwell, who in 1859 derived the momentum distribution later named after him. Boltzmann subsequently generalized Maxwell's distribution for the case of gases in external fields.

F. History of Equilibrium Statistical Mechanics

In association with Clausius, in 1871, the Scottish mathematician and physicist James Clerk Maxwell formulated a new branch of thermodynamics called Statistical Thermodynamics, which functions to analyze large numbers of particles at equilibrium, i.e., systems where no macroscopic changes are occurring, such that only their average properties such as temperature T, pressure P, and volume V become precise. Soon thereafter, in 1875, the Austrian physicist Ludwig Boltzmann formulated a precise connection between entropy S and molecular motion:

$$S = k_B log(\Omega(E)) \tag{5}$$

being defined in terms of the number of possible states or degeneracy $\Omega(E)$, at fixed energy E where k_B is Boltzmann's constant.

The year 1876 was a seminal point in the development of the modern theory of statistical mechanics as the chemical engineer J. Willard Gibbs (1839-1903), the first person in America to be awarded a PhD in engineering (Yale), published an obscure 300-page paper titled: "On the Equilibrium of Heterogeneous Substances", wherein he began the formulation of the basics of partition functions, free energies and the foundations of equilibrium statistical mechanics. The term Statistical Mechanics was introduced by Gibbs at a meeting of the American Association for the Advancement of Science in Philadelphia in 1884. This was one of the rare occasions when Gibbs went to a meeting away from New Haven. He had been professor of mathematical physics at Yale University since 1871 and had served nine years without salary. Only in 1880, when he was on the verge of accepting a professoship at John Hopkins University, did his institution offer him a salary. He had realized that the papers of Maxwell and Boltzmann initiated a new discipline which could be applied to bodies of arbitrary complexity moving according to the laws of mechanics which were investigated statistically. In the years following 1884 he formulated a general framework for Statistical Mechanics and in 1902 published his treatise that lays out many of the foundations of free energies and partition functions. Gibbs started his consideration with the principle of conservation of the phase space volume occupied by a statistical ensemble of mechanical systems. He considered three types of ensembles. The so-called microcanonical ensemble of Gibbs is an ensemble of isolated systems which all have the same energy. Boltzmann called this ensemble "Ergoden". In this case each member of the ensemble corresponds to a different microstate and all microstates have the same probability. The canonical ensemble of Gibbs corresponds to systems in contact with a heat bath. In this case the energy of the individual systems is allowed to fluctuate around the mean value E. If E_ν is the energy of an individual system ν of the ensemble, its probability p_ν is proportional to an exponential function linear in the energy $p_\nu \propto Exp(-E_\nu/k_BT)$ which is nowadays often called the Boltzmann factor. For the grand-canonical ensemble of Gibbs not only the energy but also the number of particles N_ν of the individual systems is allowed to fluctuate around the mean value N.

At equilibrium a system does not prefer any of its available equi-energetic microstates. Given $\Omega(E)$ microstates at a particular energy, the probability of finding the system in a particular microstate is simply $p = 1/\Omega(E)$. Of course physical systems are always evolving dynamically at the microscale so the assumption above assumes that regardless of the detailed dynamics of the problem, the same long-time probability occurs for each microstate that has the same energy. This is highly non-trivial and is actually completely wrong for many non-interacting systems! However once interactions are included it turns out to be true in an important but restricted sense. The validity of the hypothesis depends on the time scale of the experiments and on the structure of the energy landscape of the system under consideration.

The postulate is justified in part, for classical systems, by Liouville's theorem, which shows that if the distribution of system points through accessible phase space is uniform at some time, it remains so at later times. Similar justification for a discrete system is provided by the principle of detailed balance. Thus, if the dynamics of a system leads to an equal probability of occupancy of all microstates at the same energy, then the system is "well mixed", it exhibits"molecular chaos", it is called "ergodic", and at long times Equilibrium Statistical Physics applies. However the time scale required to achieve complete mixing can be impossible in some cases, such as glasses, in which case hysteresis or "memory" effects occur. In fact on experimental times many systems are not ergodic. In the modern literature there is much discussion of "complex energy landscapes" where non-ergodicity is typical. Nevertheless non-equilibrium extensions of statistical physics and thermodynamics can be applied to understand many aspect of even these non-ergodic cases.

II. FOUNDATIONS OF STATISTICAL MECHANICS

The infinite monkey theorem states that a monkey hitting keys at random on a typewriter keyboard for an infinite amount of time will almost surely type a given text, such as the complete works of William Shakespeare.

Now it is time to state, understand and then justify the following four laws that form the foundations of Equilibrium Statistical Physics (ESP). Keep in mind that statistical mechanics applies to the behavior of very large numbers of interacting subsystems, with some examples being spin systems, atomic systems, Fermion and Bose systems, quantum fields etc.

Axiom (i) The fundamental postulate of equilibrium statistical mechanics:
At equilibrium all accessible microstates with the same energy are found with equal probability.
A mathematical expression of this statement is as follows. If $\Omega(E)$ is the number of accessible microstates at energy E, the probability of being in any one state is $p(E) = 1/\Omega(E)$.
Corollary 1: A closed system at equilibrium is in a state of maximum entropy.
Corollary 2: Equilibrium dynamics is ergodic, so that averages over sufficiently long times and ensemble averages give the same result.

Axiom (ii) The definitions of entropy
The Boltzmann/Planck Entropy is given by: $S = k_B ln(\Omega(E,N,V))$**, where** k_B **is Boltzmann's constant**.
Corollary 1: Entropy is also given by the Gibb's formula, $S = -k_B \sum_i p_i ln(p_i)$, where p_i is the probability of a state with energy E_i.
Corollary 2: In quantum systems entropy is given by the von Neumann formula, $S = -k_B tr(\hat{\rho} ln(\hat{\rho}))$, where $\hat{\rho}$ is the density matrix, $\hat{\rho} = e^{-\beta\hat{H}}/Z$ with $\hat{H}$ the Hamiltonian and $Z = tr(e^{-\beta\hat{H}})$.

Axiom (iii) The fundamental thermodynamic relation
The fundamental thermodynamic relation is: $dU = dQ - dW = TdS - PdV + \mu dN + ...$**,**
where μ is the chemical potential of one species.

Axiom (iv) The second law of thermodynamics (Boltzmann H-theorem)
In a closed system, entropy either increases or stays the same, $\frac{dS}{dt} \geq 0$**.**

A. Arguments supporting axiom (i)

Perhaps the most general argument in favor axiom (i) is then through the master equation that focuses on the transition rates between states. Each state of a many body system has probability p_i, and the transition rate between this state and another state, with probability p_j, is given by ω_{ij}. The rate of change of the probability p_i is given by the deceptively simple but very important "master equation"

$$\frac{dp_i}{dt} = \sum_j (\omega_{ji} p_j - \omega_{ij} p_i). \tag{6}$$

At equilibrium the probability p_i is independent of time, so that,

$$\sum_j (\omega_{ji} p_j - \omega_{ij} p_i) = 0 \quad \text{at equilibrium.} \tag{7}$$

Now note that we expect the forward rate ω_{ij}, and reverse rate ω_{ji} to be the same at equilibrium in a closed system, which immediately implies that $p_i = p_j$ as required by axiom (i). This rate theory applies to both classical and quantum systems, though of course the rates are different and in the purely quantum case are often calculated using tunneling theory. In the classical case, the rates are often taken to be thermally activated.

Another general result supporting axiom (i) is Liouville's theorem for the evolution of phase space density of a classical system. In a classical system with N particles, a state of the system is described by a set of positions $\vec{q_i}$ and a set of momenta $\vec{p_i}$. The probability of finding the system in a state with a given set of positions and momenta is given by the phase space density, $\rho(\{\vec{q_i}\}, \{\vec{p_i}\})$. Liouville's theorem states that the time evolution of the phase space density is given by,

$$\frac{\partial \rho}{\partial t} + \sum_{i}^{3n} [\dot{q}_i \frac{\partial(\rho)}{\partial q_i} + \dot{p}_i \frac{\partial(\rho)}{\partial p_i}] = 0 \tag{8}$$

This is often written in the compact form,

$$\frac{\partial \rho}{\partial t} + \{\rho, H\} = 0 \tag{9}$$

Actually, Liouville's equation is a conservation equation as it follows from the fact that the Hamiltonian dynamics conserves the phase space volume because there is no dissipation. Therefore ρ obeys the continuity equation in the $6N$ dimensional phase space,

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \nabla \cdot (\vec{j}) = \frac{\partial \rho}{\partial t} + \sum_{i}^{3n} [\frac{\partial(\rho \dot{q}_i)}{\partial q_i} + \frac{\partial(\rho \dot{p}_i)}{\partial p_i}] = 0 \tag{10}$$

This reduces to the Liouville equation by noting that,

$$\rho \sum_i [\frac{\partial(\dot{q}_i)}{\partial q_i} + \frac{\partial(\dot{p}_i)}{\partial p_i}] = \rho \sum_i [\frac{\partial^2 H}{\partial q_i \partial p_i} - \frac{\partial^2 H}{\partial p_i \partial q_i}] = 0 \tag{11}$$

where we used Hamilton's equations, $\dot{p} = -\partial H/\partial q$, $\dot{q} = \partial H/\partial p$. Liouville's equation is just the equation for a conserved current of the probability density in phase space, so that if the initial phase space density is a constant everywhere, it remains constant for all time. However, it is not true that a constant density is achieved for any starting condition for any density, that is, the uniform state is not necessary stable to small random perturbations.

The density matrix $\hat{\rho} = exp(-\beta \hat{H})/Z$, where Z is the canonical partition function and is time independent, is the quantum mechanical extension of the classical phase space density. $\hat{H}$ is the Hamiltonian. von Neumann showed that the density matrix has the following time evolution,

$$i\hbar \frac{\partial \hat{\rho}}{\partial t} + [\hat{\rho}, \hat{H}] = 0. \tag{12}$$

It is relatively easy to derive this from the Schrödinger equation but it is more difficult to take the classical limit and show that it reduces to the Liouville equation.

Despite the fact that the uniform solution is consistent with the Master equation and the Liouville equation, axiom (i) is actually counter-intuitive as many non-interacting systems do not occupy all states with equal probability. There are (at least) two major issues regarding the validity of axiom (i):

Firstly does the dynamics explore all of the available phase space? In fact non-interacting systems and most of the interacting systems we study in physics courses are not ergodic. For example consider a system of coupled harmonic oscillators. In these systems if you put the system in one normal mode it stays there for all time. There has been intense study of the conditions for the "onset of chaos" in dynamical systems. In low dimensional systems interesting routes to chaos have been discovered and explored computationally. However, "integrable systems", that is those whose dynamics we can solve for analytically are non-ergodic. This means that to really study chaotic systems we need to use computational approaches. The consensus after much head scratching, analysis and simulations is that interactions that produce chaotic dynamics are essential to the validity of axiom (i). However in systems with many degrees of freedom even weak interactions lead to mixing of states so that at long times all accessible states mix in accordance with axiom (i). In terms of normal modes, the introduction of interactions leads to transitions between different states or normal modes of the many body system so that at long times all modes with equal energy have the same probability of occupancy.

The second key issue regarding axiom (i) addresses the issue of the timescales required to achieve ergodic dynamics. The slowest timescales of the dynamics are the most important for this issue, and in the Master equation of Eq. (6) lead to the smallest transition rates. If one of the transition rates in this equation is longer than the experimental time, then the experiment is in the non-ergodic regime. There are many systems that are non-ergodic due to large energy barriers that separate regions of phase space into distinct subdomains. Perhaps the simplest example is a magnetic storage bit on a hard drive disk. These days each bit is around $10nm^3$ and contains thousands of atomic scale spins that are aligned in the same direction. Though the two possible orientations of the magnetic domain or bit are equal in energy, the dynamics does not explore both orientations at room temperature, so the system is non-ergodic. In general at low enough temperatures, all many particle systems are non-ergodic! Once there is a phase transition, the phase space breaks up into separated regions. The dynamics on experimental timescales only explores one region, depending on the initial condition of the experiment. This behavior is nicely illustrated by simulations. This behavior is often represented pictorially by drawing energy landscapes. In the case of the magnetic system there are two minima in the energy at low temperature, with a large barrier in between. Protein folding is believed to have a funnel landscape, that has a complex and rough component. Many frustrated systems, such as glasses have complex energy landscapes with barriers on a wide range of energy scales. This leads to hysteresis, aging and memory effects in most experiments on these systems.

B. Arguments supporting axiom (ii)

Even the early work on thermodynamics of engines noted the relation between the “freedom” of a working gas and the “heat” content of the gas. The first law of thermodynamics makes this explicit by relating the change in internal energy to the heat added to a system minus the work done by the system $\Delta U = Q - W$. Clausius introduced the entropy through $dS = dQ/T$, so the entropy must be an extensive quantity and TS must have the units of energy, and it must be extensive. Boltzmann proposed his form of the entropy to have these properties.

In order to understand the “freedom” or degeneracy in a system, we first consider non-interacting systems. A simple example is a spin half model of magnetism where each spin can be in one of two states “spin up” or “spin down”. If we have N spins, the total number of available states is $\Omega(E, N) = 2^N$. In a similar way, a classical gas has degeneracy $\Omega(E, N) \propto V^N$. To find a quantity that is extensive we must take a logarithm of these degeneracies, so that $S \propto ln(\Omega(E))$.

It is easy to show that Corollary 1 of axiom (ii) is consistent with the Boltzmann entropy, by using $p_i = 1/\Omega(E)$, so that,

$$S = -k_B \sum_i p_i ln(p_i)) = -k_B \Omega(E) \frac{1}{\Omega(E)} ln(\frac{1}{\Omega(E)}) = k_B ln(\Omega(E)) \tag{13}$$

By transforming to the basis of energy eigenstates, the von Neumann entropy of Corollary (ii) reduces to the entropy of Corollary (i),

$$S = -k_B tr(\hat{\rho} ln(\hat{\rho}) = -k_B tr(U^\dagger \hat{\rho} U U^\dagger ln(\hat{\rho}) U) \tag{14}$$

where U is the unitary matrix that diagonalizes both the Hamiltonian $\hat{H}$ and the density matrix $\hat{\rho} = Exp(-\beta \hat{H})/Z$. These matrices can be introduced into the trace formula as the trace formula is invariant under cyclic permutations of a matrix product, and the unitary matrix has the property $UU^\dagger = unit\ matrix$. In the (many body) eigenfunction basis, the von Neumann entropy is then,

$$S = -k_B tr(U^\dagger \hat{\rho} ln(\hat{\rho}) U) = -k_B \sum_i p_i ln(p_i) \tag{15}$$

where, in the canonical ensemble, $p_i = exp(-\beta E_i)/Z$, where $Z = \sum_i e^{-\beta E_i}$. In carrying out this reduction, we also use the fact that $\hat{\rho}$ and $ln(\hat{\rho})$ commute and hence are diagonalized by the same unitary matrix.

C. Arguments supporting axiom (iii)

Axiom (iii) follows from the first law of thermodynamics, the Clausius relation between heat and entropy, with the addition of the contribution due to changing the particle number. The first law of thermodynamics states $\Delta U(S, V) = Q - W \to TdS - PdV$. If we include the possibility of changing the number of particles in the system, the change in internal energy is μdN, where μ is the energy required to add a particle to the system. Note that there are a variety of definitions of μ and the issue is subtle in systems with energy gaps, such as semiconductors. As long as we clearly define the process underlying our definition of μ there is no ambiguity. Here, our definition is adding a particle to the system, though we could also define it through removing a particle, or by taking an average of the addition and removal processes. In any case once we add the energy due to adding a small number of particles dN, with μ constant, we find,

$$dU(S, V, N) = TdS - PdV + \mu dN \tag{16}$$

as stated in axiom (iii). This will form the basis of many of the derivations of thermodynamic relations that we need to carry out, so it is worth memorizing.

It is important to note that the axiom (iii) may be stated much more generally, as $dU = dQ - dW$, with $dQ = TdS$ as above, but with the possibility of much more general forms of the "generalized" work. A generalized work expression that includes magnetostatic and electrostatic terms as well as many chemical species is,

$$dW = PdV - \vec{H} \cdot d\vec{M} - \vec{E} \cdot d\vec{P} - \sum_i \mu_i dN_i + ... \tag{17}$$

where $\vec{H}$ and $\vec{E}$ are the applied magnetic and electric fields, while $\vec{M}$ and $\vec{P}$ are the magnetization and polarization.

These, and other terms are added to treat problems arising in magnetic, superconducting and dielectric materials, as well as systems with many components. Further generalizations may be required if the incremental work is more complex or anisotropic. For most of the course we shall focus upon the special case where the mechanical work is isotropic so we get PdV and the system is monatomic so the chemical potential term is μdN. In chemistry there must be more than one species so that chemical reactions can occur so in those applications the term $\sum_i \mu_i dN_i$ is required. Chemists often want to determine the amount of heat released or absorbed in a reaction, i.e. is the reaction endothermic (heat absorbtion) or exothermic (heat released), and the equilibrium constants of the reaction, i.e. at chemical equilibrium what is the concentration of reactants and products.

D. Arguments supporting axiom (iv) - Boltzmann's H theorem

For a broad range of systems axiom (iv) can be proven by starting with a time derivative of the Gibb's entropy,

$$\frac{\partial S}{\partial t} = -k_B \sum_i [ln(p_i)\frac{dp_i}{dt} + \frac{dp_i}{dt}] = -k_B \sum_i ln(p_i)\frac{dp_i}{dt} \tag{18}$$

where in the last equation we used the fact that $\sum_i p_i = 1$ so that $\sum_i dp_i/dt = 0$. Now we can use the Master equation (with $\omega_{ji} = \omega_{ij}$) to write,

$$\frac{dS}{dt} = -k_B \sum_{i \neq j} \omega_{ij}(p_j - p_i)log(p_i) \quad \text{or} \quad \frac{dS}{dt} = -k_B \sum_{i \neq j} \omega_{ij}(p_i - p_j)log(p_j) \tag{19}$$

where we exchanged the indices in the last expression on the RHS. Summing the two forms and dividing by two, we have,

$$\frac{dS}{dt} = -\frac{1}{2}k_B \sum_{i \neq j} \omega_{ij}(p_i - p_j)(log(p_j) - ln(p_i)) = \frac{1}{2}k_B \sum_{i \neq j} \omega_{ij}(p_i - p_j)[log(p_i/p_j)] \tag{20}$$

Now notice that the two terms in brackets ()[] are either both positive (if $p_i > p_j$) or both negative (if $p_i < p_j$), so that the right hand side of this equation is either positive or zero, as stated in axiom (iv).

III. EQUILIBRIUM CONDITIONS, PARTITION FUNCTIONS AND FREE ENERGIES

Mathematics is a language - J. Willard Gibbs

A. Equilbrium conditions

Consider two systems, systems 1 and 2, characterized by their entropy, volume and number of particles, S_1, V_1, N_1) and S_2, V_2, N_2 respectively. The two systems are at equilibrium with each other, but they do not affect each other in any other way. They are in contact and can exchange both energy and particles. The combined system has entropy,

$S = S_1 + S_2$, volume $V = V_1 + V_2$, and number of particles $N = N_1 + N_2$. The combined system is isolated, and is at equilibrium. No work is done by either system as the volumes V_1 and V_2 do not change. The change in internal energy is then only due to exchange of heat and particles. The density of states of the systems are $\Omega_1(E_1, N_1), \Omega_2(E_2, N_2)$. We then also have $E = E_1 + E_2$. With this construction, we now consider the fundamental thermodynamic relation,

$$dU_1 = dE_1 = T_1 dS_1 - P_1 dV_1 + \mu_1 dN_1, \quad \text{and} \quad dU_2 = dE_2 = T_2 dS_2 - P_2 dV_2 + \mu_2 dN_2 \tag{21}$$

In the combined system E, S, V, N are fixed, so we can write,

$$dU_2 = dE_2 = T_2 dS_2 - P_2 dV_2 + \mu_2 dN_2 = d(E - E_1) = T_2 d(S - S_1) - P_2 d(V - V_1) + \mu_2 d(N - N_1) \tag{22}$$

which reduces to

$$-dE_1 = -(T_2 dS_1 - P_2 dV_1 + \mu_2 dN_1) \tag{23}$$

Comparing the first of Eq. (21) with Eq. (23) we find the equilibrium conditions,

$$T_1 = T_2; \quad P_1 = P_2; \quad \mu_1 = \mu_2 \tag{24}$$

The analysis above obscures an important approximation that is made in the analysis. This approximation is revealed by considering instead an analysis based on the density of states. The density of states of the combined system is,

$$\Omega(E, E_1, E_2) = \Omega(E_1)\Omega(E_2) \ \text{ with } \ E = E_1 + E_2, N_1, N_2 \text{ fixed.} \tag{25}$$

so that,

$$\Omega(E) = \sum_{E_i} \Omega_1(E_i)\Omega_2(E - E_i) = \sum_{E_i} \Omega_2(E_i)\Omega_1(E - E_i) \tag{26}$$

where we have to sum over the available states of one of the subsystems as we have only one constraint on the total energy. However these functions are extremely rapidly varying with energy, so there is an extremely sharp maximum in the total density of states at some value of E_i, that we take to be E_m. We therefore have,

$$\Omega(E) \approx \Omega_1(E_m)\Omega_2(E - E_m) \approx \Omega_2(E_m)\Omega_1(E - E_m) \tag{27}$$

Since the total energy E is a constant, we have,

$$\frac{\partial \Omega}{\partial E_m} = 0 = \frac{\partial \Omega_1(E_m)}{\partial E_m}\Omega_2(E - E_m) - \Omega_1(E_m)\frac{\partial \Omega_2(E - E_m)}{\partial E_m}. \tag{28}$$

Dividing by $\Omega_1(E_m)\Omega_2(E - E_m)$, we find,

$$\frac{1}{\Omega_1(E_m)}\frac{\partial \Omega_1(E_m)}{\partial E_m} = \frac{1}{\Omega_2(E - E_m)}\frac{\partial \Omega_2(E - E_m)}{\partial E_m} \tag{29}$$

which may be written as,

$$\left(\frac{\partial S_1}{\partial U_1}\right)_{V_1, N_1} = \left(\frac{\partial S_2}{\partial U_2}\right)_{V_2, N_2} \tag{30}$$

The conditions for equilibrium are constant temperature, constant pressure and equal chemical potential for each species in the system.

B. Boltzmann weights and partition functions

There are several ways to show that the Boltzmann weight is correct to describe the way in which the importance of a state varies with its energy. All of them are based on the fundamental postulate of equilibrium statistical mechanics. One approach is to consider that one of the subsystems, is small compared to the other. Lets take subsystem 1 to be small and subsystem 2 to be large. In this case, we call subsystem 2 the reservoir. Then of course $E_2 >> E_1$. The reservoir is at temperature T. Consider the case where no particle exchange occurs, but heat exchange does.

The probability a state with energy ϵ in system 1 is occupied is $p_1(\epsilon) = 1/\Omega_1(\epsilon)$. Since $\Omega(E) = \Omega_1(E_1)\Omega_2(E_2)$, $p(E) = p_1(E_1)p_2(E_2)$. Now the energy of the total system is conserved so that $p(E)$ is a constant.

$$p_1(E_1) = p(E)\Omega_2(E-E_1) = const\ exp(S_2(E-E_1)/k_B) \approx const\ exp[S_2(E) - \frac{E_1}{k_B}\frac{\partial S_2}{\partial E_2} + ...] \rightarrow Aexp(-\beta E_1) \tag{31}$$

This is of course the Boltzmann factor in the cannonical ensemble. We normalize the Boltzmann factor to make it a probability, by setting $A = 1/Z$, where Z is the canonical partition function.

A second approach is the use the maximum entropy principle, so that if the entropy is given by the Gibb's form,

$$S = -k_B \sum_i p_i ln(p_i) \tag{32}$$

then the probabilities will try to maximize the entropy. In the microcanonical ensemble, this maximization must preserve normalization of the probabilities so that $\sum_i p_i = 1$. Imposing this constraint, we have,

$$S = -k_B \sum_{i=1}^{M-1} p_i ln(p_i) - k_B(1 - \sum_{i=1}^{M-1} p_i) ln(1 - \sum_{i=1}^{M-1} p_i) \tag{33}$$

Taking a derivative with respect to p_i yields,

$$\frac{\partial S}{\partial p_i} = -k_B[ln(p_i) - ln(1 - \sum_{i=1}^{M-1} p_i)] = 0 \quad \text{or} \quad p_i = 1 - \sum_{i=1}^{M-1} p_i \tag{34}$$

This is solved by $p_i = 1/M$ where M is the number of states at energy E (i.e. $M = \Omega(E)$). This calculation can also be carried out using a Lagrange multiplier where we introduce the constaint by writing,

$$S = -k_B \sum_i p_i ln(p_i) - \lambda(\sum_i p_i - 1) \tag{35}$$

The entropy is then maximized with respect to p_i and with respect to the Lagrange multiplier λ. The latter condition leads to the constraint that we want to impose, ie. $\sum_i p_i = 1$. Setting the derivative with respect to p_i to zero now gives,

$$p_i = e^{-k_B - \lambda} \tag{36}$$

This is a constant, just as we found before. However we have to fix the constant using other information, i.e. $e^{-k_B-\lambda} = 1/M$. The Lagrange multiplier procedure is easily extended to the cannonical and grand-canonical distribution functions by imposing further conditions using additional Lagrange multipliers.

In the canonical ensemble the energy is not fixed, but we have to ensure that the relation $\sum_i p_i \epsilon_i = E$ so that the correct average energy is preserved. To impose that condition, we consider a second Lagrange multiplier, β,

$$S = -k_B \sum_i p_i ln(p_i) - \lambda(\sum_i p_i - 1) - k_B\beta(\sum_i p_i\epsilon_i - E) \tag{37}$$

Taking a variation of this with respect to p_i and setting the resulting equation to zero yields,

$$p_i = e^{-k_B - \lambda - \beta\epsilon_i} = \frac{e^{-\beta\epsilon_i}}{Z} \tag{38}$$

where

$$e^{-k_B-\lambda} = \frac{1}{Z} \quad \text{and} \quad Z = \sum_i e^{-\beta\epsilon_i} \tag{39}$$

This procedure generalizes easily to any ensemble, for example if we consider the grand canonical ensemble where $\sum_i p_i n_i - N$ is the number of particles, then the entropy is

$$S = -k_B \sum_i p_i ln(p_i) - \lambda(\sum_i p_i - 1) - k_B\beta(\sum_i p_i\epsilon_i - E) + k_B\beta\mu(\sum_i p_i N_i - N) \tag{40}$$

so the probability is,

$$p_i = e^{-k_B - \lambda - \beta\epsilon_i + \beta\mu N_i} = \frac{e^{-\beta(\epsilon_i - \mu N_i)}}{\Xi} \tag{41}$$

where $\Xi = \sum_i e^{-\beta(\epsilon_i - \mu N_i)}$ is the grand partition function.

C. Quantum systems

So far we have talked about phase space where the density in phase space obeys the Liouville equation and we have talked a lot about systems where we have a set of energy levels. In quantum systems we also have a set of energy levels, so the formalism involving energy levels is appropriate. Everything we have discussed so far thus applies to the quantum case, provided we work in the basis of the eigenfunctions of the Hamiltonian operator, and provided we take expectations of operators that commute with the Hamiltonian. To treat more general operators, we need to extend the formulation. The most straightforward way to carry out is to generalize the partition function in the canonical ensemble, where,

$$Z = \sum_i e^{-\beta E_i} = tr(\lambda) \tag{42}$$

where λ is a matrix with Boltzmann weights on its diagonal. Now if the Hamiltonian is $\hat{H}$ and is diagonalized by a unitary transformation U, which has the properties $U^\dagger U = UU^\dagger = Unit\ matrix$ so that,

$$U^\dagger \hat{H} U = \lambda_H, \quad \text{then} \quad \hat{H} = U \lambda_H U^\dagger \tag{43}$$

where λ is a diagonal matrix with the eigenvalues of $\hat{H}$ on the diagonal, so that,

$$tr(U\hat{H}U^\dagger) = tr(\hat{H}) \tag{44}$$

where we used the cyclic property of the trace, ie. $tr(ABC) = tr(BCA)$. Now we note that the matrix $e^{-\beta\hat{H}}$ is diagonalized by the same matrix as $\hat{H}$, which follows from expanding the exponential, therefore,

$$tr(Ue^{-\beta\hat{H}}U^\dagger) = tr(\lambda) = tr(e^{-\beta\hat{H}}). \tag{45}$$

We thus find that in any quantum mechanical basis, the cannonical partition function is given by,

$$Z = tr(e^{-\beta\hat{H}}) \tag{46}$$

It is then natural to define the density matrix in the cannonical ensemble as,

$$\hat{\rho} = \frac{e^{-\beta\hat{H}}}{Z} \tag{47}$$

From which it is easy to show that all operators that commute with $\hat{H}$ reduce to the expected form in the diagonal representation. It is a little more work to show that ensemble averages of operators that do not commute with $\hat{H}$ are also correctly desribed by this expression. Note that in some books the density matrix is defined as $e^{-\beta\hat{H}}$ while in other it is defined as $e^{-\beta\hat{H}}/Z$. Here we use the latter definition.

The generalization of the grand partition function follows in a similar manner so that,

$$\Xi = tr(e^{-\beta(\hat{H}-\mu\hat{N})}) \tag{48}$$

where $\hat{N}$ is the number operator.

D. Relation between partition functions and entropy

These relations are found by using the Gibbs formula for the entropy,

$$S = -k_B \sum_i p_i ln(p_i). \tag{49}$$

Using the Boltzmann probability in the canonical ensemble $p_i = exp(-\beta\epsilon_i)/Z$, we have,

$$S = -k_B \sum_i p_i [\frac{-\epsilon_i}{k_B T} - ln(Z)] = \frac{U}{T} + k_B ln(Z) \quad \text{so} \quad U - TS = -k_B T ln(Z) \tag{50}$$

Using the Boltzmann probability in the grand canonical ensemble $p_i = exp[-\beta(\epsilon_i - \mu N_i)]/\Xi$, we have,

$$S = -k_B \sum_i p_i[\frac{-(\epsilon_i - \mu N_i)}{k_B T} - ln(\Xi)] = \frac{U - \mu N}{T} + k_B ln(\Xi) \quad \text{so} \quad U - TS - \mu N = -k_B T ln(\Xi) \tag{51}$$

Because of these relations, we define two new free energiess, the Helmholz energy $F(or\ A) = U - TS$ and the Grand Potential $\Phi_G = U - TS - \mu N$, so that we have the fundamental relations between the partition functions and free energies,

$$F = U - TS = -k_B T ln(Z); \quad \text{and} \quad \Phi_G = U - TS - \mu N = -k_B T ln(\Xi) \tag{52}$$

E. Interesting Aside: Connection with information theory - Shannon entropy

Consider information transmission down a noisy communication channel. If a signal contains N bits, and we want to transmit a signal that has a fix value for each bit, then the perfect transmission case corresponds to $p_i = 1$, where p_i is the probability of the perfect signal state amongst the 2^N states possible with N binary bits. Shanon defined the information function (really its the ignorance function!)

$$I = -\sum_i \rho_i log_b \rho_i = \langle -log_b \rho \rangle. \tag{53}$$

When all the probabilities (ρ_i) are equal, I is maximal, and we have minimal information about the system. When our information content is maximal (i.e., one rho is equal to one and the rest are zero, such that we know what state the system is in), the information function is minimal. The logarithm can to be to any base b, though for digital electronics it is usually 2. When the base is e, the information function is the thermodynamic entropy, provided we multiply by $-k_B$. Maximum information transmission corresponds to minimization of the Shannon entropy.

F. Derivation of two key thermodynamic relations: $U = TS - PV + \mu N$; and $SdT - VdP + Nd\mu = 0$.

First note that $U - TS + PV = \mu N$ is NOT the same as the fundamental thermodynamic relation $dU = TdS - PdV + \mu dN$.

To discover $U - TS + PV = \mu N$ note that there is a general mathematical theory of homogeneous functions and homogeneous equations, where a homogeneous function has the property that,

$$f(\lambda x, \lambda y, \lambda z) = \lambda^k f(x, y, z). \tag{54}$$

Because the energy (or entropy) are defined in terms of extensive variables and are themselves extensive, they are also homogeneous, so that,

$$\lambda U(S, V, N) = U(\lambda S, \lambda V, \lambda N) \tag{55}$$

which is a linear homogeneous relation ($k = 1$). Now we take a derivative of this relation with respect to λ and using the chain rule we find,

$$U(S, V, N) = \left(\frac{\partial U}{\partial(\lambda S)}\right)_{V,N} S + \left(\frac{\partial U}{\partial(\lambda V)}\right)_{S,N} V + \left(\frac{\partial U}{\partial(\lambda N)}\right)_{S,V} N \tag{56}$$

This equation is true for all λ. If we take $\lambda = 1$ we have,

$$U = \left(\frac{\partial U}{\partial S}\right)_{V,N} S + \left(\frac{\partial U}{\partial V}\right)_{S,N} V + \left(\frac{\partial U}{\partial N}\right)_{S,V} N \tag{57}$$

Howerver, from the fundamental thermodynamic relation we also have,

$$dU = TdS - PdV + \mu dN = \left(\frac{\partial U}{\partial S}\right)_{V,N} dS + \left(\frac{\partial U}{\partial V}\right)_{S,N} dV + \left(\frac{\partial U}{\partial N}\right)_{S,V} dN \tag{58}$$

so that,

$$T = \left(\frac{\partial U}{\partial S}\right)_{V,N} \quad ; \quad P = -\left(\frac{\partial U}{\partial V}\right)_{S,N} \quad ; \quad \mu = \left(\frac{\partial U}{\partial N}\right)_{S,V}. \tag{59}$$

Using these relations in Eq. (57), we find

$$U = TS - PV + \mu N \tag{60}$$

which is the first relation that we seek. To find the Gibbs-Duhem relation, we note that now we have two expressions for dU, so that

$$dU = d(TS - PV + \mu N) = TdS - PdV + \mu dN \tag{61}$$

from which we deduce the Gibbs-Duhem relation,

$$SdT - VdP + Nd\mu = 0. \tag{62}$$

IV. OUTLINE OF KEY MANY BODY STATISTICAL PHYSICS SYSTEMS

Statistical physics systems may be formulated in many different ways depending on the problem. The key systems are: classical and quantum interacting particle systems, "classical" and quantum interacting spin systems, classical and quantum interacting field systems.

An example of a classical particle Hamiltonian is represented by the Hamiltonian,

$$H = \sum_{i=1}^{N} \frac{\vec{p}_i^{\,2}}{2m} + V(\{\vec{r}_i\}). \tag{63}$$

The fact that the momentum only appears quadratically leads to the equipartition theorem for the kinetic energy. An important special case is the case of pair potentials, such as gravity, coulomb and Lennard-Jones interactions where $V_N(\{\vec{r}_i\}) = \sum_{j>i}^{N} V_2(\vec{r}_{ij})$. The classical particle Hamiltonian is used for simulations in all areas of science, including simulations of gravitating systems, simulations of proteins and polymers and simulations of atomic solid state systems. Simulations are often carried out using Molecular dynamics, a method we shall talk in detail about later. In complex systems such as proteins and solid state materials developing good potentials is a large part of the problem and these developments are often based on ab-intio quantum methods. The simulation procedures however rely on statistical physics and thermodynamics. An example of a quantum particle Hamiltonian (or many body Hamiltonian) is the Hubbard model,

$$H = \sum_{j>i} t_{ij} c_i^{\dagger} c_j + U \sum_i n_{i,\uparrow} n_{i,\downarrow} \tag{64}$$

where c_i is the creation operator, and n_i is the number operator. This is a first approximation to the description of electrons in materials with strong Coulomb interactions. The first term describes the band energy or kinetic energy, while the second term describes the on-site Coulomb repulsion.

An example of a classical or quantum spin system is,

$$H = \sum_{j>i} J_{ij} \vec{S}_i \cdot \vec{S}_j. \tag{65}$$

In the simplest the spin is a scalar that has two values, up or down. This is the Ising spin case. More general scalar spin models allow the spin to take values $-S_{max}, -S_{max}+1, ...S_{max}$. Vector spin models often consider the vecor $\vec{S}$ to have n components and to be normalized to magnitude one. In that case the $n = 1$ case corresponds to a vector that can have values on the perimeter of a circle, so it is called the $O(1)$ model. This applies for example to vortex states in superconductivity. If the spin vector has n components it is called an $O(n)$ model as the vector lies on the surface of a sphere in $n - dimensional$ space. In classical spin models the spins operators are classical variables, while in quantum spin models they are quantum operators. For example the Hamiltonian () for the case of a spin 1/2 quantum system with O(3) symmetry is the Heisenberg Hamiltonian,

$$H = \sum_{j>i}^{N} J_{ij} \vec{S}_i \cdot \vec{S}_j = \sum_{j>i}^{N} J_{ij} (\sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y + \sigma_i^z \sigma_j^z) \tag{66}$$

where $\sigma^x, \sigma^y, \sigma^z$ are the Pauli matrices.

Field theories do not have an underlying lattice and are instead described in terms of fields at each point in space. This is familiar from electricity and magnetism which is a classical field theory, and quantum electrodynamics which is a quantum field theory. The methods of field theory have played an important role in modern statistical physics. Landau developed a field theory to describe the phase behavior of statistical systems based on very general symmetry arguments. Ginzburg-Landau theory adds fluctuations to this analysis and this theory is the most successful for treating the phase behavior of complex many body systems. We shall discuss Landau theory and the Ginzburg-Landau approach in the third section of the course. One important example is the Landau field theory for the superconducting or superfluid state in the absence of a field,

$$H = \int [-\frac{\hbar^2}{2m}|\nabla\psi|^2 + \alpha|\psi|^2 + \beta|\psi|^4] d^3r \tag{67}$$

where ψ is a complex field (wavefunction). A similar model may be written down as a continuum limit for spin models, such as the Ising model. Using a field theory approach to the magnetic phase transition Ken Wilson applied quantum field theory methods to statistical physics leading to analytic calculation of critical exponents characterizing behavior near second order phase transitions. This work was acknowledged with a nobel prize in 1982. Later we shall extend this Hamiltonian to treat the effect of an applied magnetic field in superconductors and charged superfluids, enabling effective models of the mixed and intermediate states of superconductors in a field. Alexei Abrikosov used this field theory to theoretically demonstrate the existence of the flux lattice in superconductors and his work was acknowledged with a nobel prize in physics in 2003. We can also add an external potential to describe the behavior of superfluid in an atomic trap. Carl Wieman and Eric Cornell developed experimental methods to demonstrate Bose condensation of Rubidium atoms in an optical and magnetic trap. Their work was rewarded with the 2001 Nobel prize.

V. COMPUTATIONAL METHODS

These days simulation studies in statistical physics are as important as analytic analysis so that every physicist needs to know the basics of computational methods and the strengths and weaknesses of each approach. The most important computational methods in statistical physics can be placed in two broad categories, Monte Carlo (MC) methods and Molecular Dynamics (MD) methods. There are classical and quantum versions of both of these methods. Here we discuss classical MC and classical MD methods applied to classical spin systems and classical particle systems respectively. MC and MD methods provide systematic methods to sample configurations in phase space with the correct Boltzmann probabilities for a statistical physics system. MC methods directly sample configurations, while MD methods follow trajectories in phase space using a physically correct particle dynamics. First we discuss the MC method using the Ising model as an example.

A. Classical Monte Carlo simulations

In statistical physics Monte Carlo methods are considered to have started in the Manhattan project (1940's) during which J. von Neumann, S. Ulam, N. Metropolis and R. Feynman used random numbers to estimate neutron scattering and absorption rates in materials. The Metropolis method which you will use in the first coding assignment originated in the paper: "Equation of state calculations by very fast computing machines", by N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, M. Teller and E. Teller, J. Chem. Phys. 21. 1087 (1953). A recent retrospective "Marshall Rosenbluth and the Metropolis algorithm", Physics of Plasmas 12, 57303 (2005) describes some of the history and makes it clear that Metropolis actually had little to do with the development of what is now called the Metropolis method. The credit should really go to Rosenbluth and Teller. A form of the Metropolis algorithm which is more efficient for polymer chains and proteins was also developed by Rosenbluth, and is called the Rosenbluth algorithm.

The Metropolis method is the most widely used Monte Carlo method. Monte Carlo methods can be considered to be improvements or variations on other random sampling methods such as importance sampling, which have a longer history. They are also the most widely used example of a computational method where we try to deduce the typical behavior of a complex system by repeated stochastic simulations. A few examples of the use of Monte Carlo methods in physics are:
(i) Thermodynamic quantities in statistical physics (e.g. Ising model magnetization)
(ii) Multiple integrals in mathematics and analysis
(iii) Reaction-diffusion problems in chemical engineering and geophysics

(iv) Polymer or protein conformations
(v) Particle showers in Nuclear and HEP

Monte Carlo methods are used to generate a sequence of configurations to represent the typical behavior of a statistical physics system at a given set of conditions, such as fixed temperature, volume, number or particles etc. We can use these configurations to find the average behavior of properties of interest, such as the internal energy. The statistical physics average of any operator in classical and quantum systems is given by,

$$< \hat{O} >= \sum_i O_i p_i = tr(\hat{O}\hat{\rho}) \tag{68}$$

where $\hat{H}$ is the Hamiltonian and $\beta = 1/(k_B T)$ with k_B Boltzmann's constant and T the temperature. In a Monte Carlo simulation, we generate configurations with the correct Boltzmann probability, so the probability factor in Eq. () is already accounted for. We therefore calculate a quantity, for example the energy, in each configuration and then take a simple average. The central problem in a Monte Carlo simulation is to find a set of configurations that sample the Boltzmann probability correctly. We use the classical Ising model as an example.

1. *An example - The Ising model*

The Ising model was invented to provide an understanding of phase transitions in ferromagnets, where it is known that mangetism is lost for $T > T_c$ where T_c, is the Curie temperature. The behavior near T_c is complicated and co-operative so that there is actually a singular behavior in many physical properties at T_c. In addition the geometry of the magnetic domains at T_c is fractal. All of these properties can be deduced by using the Monte Carlo method to find the average properties of Ising magnets. The Hamiltonian of a spin 1/2 Ising magnet is given by,

$$H = -J \sum_{<ij>} S_i S_j \tag{69}$$

where $J > 0$ is a ferromangetic exchange constant, $S_i = \pm 1$ is a spin variable, and the sum is over nearest neighbor sites on a lattice. We take the square lattice as our example, but the model can be defined on any graph structure, including complex networks. As well as the average energy, which approximates the internal energy, a second operator or quantity that is very important in magnetic materials is the magnetization, $m = (\sum_i < S_i >)/N$. We can use the Monte Carlo method to generate a sequence of configurations and then to calculate the energy and magnetization of each configuration.

2. *Generating configurations - the Metropolis method in the canonical ensemble*

Consider again the equilibrium Boltzmann probability of finding a state, i,

$$p_i = e^{-E_i/k_B T}/Z \quad \text{at equilibrium in the canonical ensemble} \tag{70}$$

We want to make sure that this probability is correctly recovered by our MC sampling procedure, at least if we run the sampling procedure for long enough. Of course in non-ergodic systems we can't reach all states, nevertheless we can find a stable limiting sampling of configurations in key regions of phase space.

We know that the Boltzmann probability has a time evolution given by the master equation,

$$\frac{dp_i(t)}{dt} = \sum_j [w_{ji} p_j(t) - w_{ij} p_i(t)] \tag{71}$$

where now the time dependence of the probabilities is made explicit. At long times the probabilities approach steady state, at least in ergodic systems. Now we can use a discrete form of the time derivative and set $\delta t = 0$, to find,

$$p_i(t+1) = p_i(t) + \sum_j [w_{ji} p_j(t) - w_{ij} p_i(t)] \tag{72}$$

If we know the transition probabilities w_{ij}, w_{ji}, then for a given initial set of the probabilities $\{p_i\}$ we can simulate the time evolution of the probabilities. However this is prohibitive computationally due to the fact that the phase space grows exponentially with the number of spins or particles in the system. The Metropolis method is an efficient method for sampling configurations with the correct probability.

To sample equilibrium configurations correctly, we have to ensure that the transition rates that we use lead to the equilibrium probabilities (). At equilibrium, the probabilities should have no more time dependence, so Eq. () becomes

$$\sum_j [w_{ji}p_j - w_{ij}p_i] = 0 \quad \text{for every i} \tag{73}$$

where p_i and p_j are now the equilibrium Boltzmann probabilities. This important balance equation defines the transition rates w that are consistent with equilibrium. A simple way to ensure that this equation holds is to demand that every term in the sum is zero, which gives the detailed balance condition,

$$\frac{w_{ji}}{w_{ij}} = \frac{p_i}{p_j} = e^{-(E_i - E_j)/k_B T} \tag{74}$$

A very important feature of this equation is that Z cancels on the RHS, so the ratio of transition rates depends only on the energy difference between two states. Calculating Z is difficult because it requires a sum over all states.

The Metropolis MC method gives a set of transition rates that satisfy detailed balance condition and in practice is efficient computationally. It is also simple to program. The MC procedure is as follows:

```
The Metropolis MC algorithm
Choose a spin Hamiltonian, graph structure and number of spins
Choose an initial state and boundary conditions to be used
Choose a procedure for MC “moves”, e.g. a single spin flip of a randomly chosen spin
Then, do
  1. Find the energy of an initial state, $E_i$.
  2. Find the energy, $E_j$, of a state where a randomly chosen MC move is carried out on the initial state.
  3. Choose a random number $X \varepsilon [0, 1]$.
  4. If $w_{ij} = e^{-\beta(E_j - E_i)} \geq X$, accept the MC move and store the new state as the initial state.
If $w_{ij} = e^{-\beta(E_j - E_i)} < X$, reject the MC move and keep the initial state
  5. Update averages. If averaged quantities are converged to required accuracy, Exit. If not return to 1.
```

To show that this procedure satisfies detailed balance, consider the two cases $E_j - E_i > 0$ and $E_j - E_i < 0$, which lead to,

$$w_{ij} = e^{-\beta(E_j - E_i)}; \quad w_{ji} = 1 \quad \text{if } E_j - E_i > 0 \tag{75}$$

and

$$w_{ij} = 1; \quad w_{ji} = e^{-\beta(E_i - E_j)} \quad \text{if } E_j - E_i < 0 \tag{76}$$

In both cases the detailed balance condition is satisfied.

There are many ways to choose different MC moves. A single spin flip is the simplest and is often used, however spin exchange is also used a lot. More sophisticated algorithms that flip clusters of spins are much more efficient near critical points. Algorithms such as the Swensden-Wang and Wolff algorithms are used extensively for this purpose. In glasses and in other systems that are not ergodic, a wide variety of different approaches have been tried. This area is still very active.

Later in the course we shall talk more extensively about phase transitions, however here take note that the nearest neighbor spin half ferromagnetic Ising model exhibits a second order phase transition in two and higher dimensions. The magnetization shows this phase transition very clearly and so is called an “order parameter” for the transition. At temperatures less than the critical or Curie temperature, the system is magnetized, while for $T > T_c$, the system is a paramagnet. The behavior near the Curie point, T_c, is of the form,

$$m \approx |T_c - T|^\beta \tag{77}$$

and is mathematically singular, in the thermodynamic limit. At the Curie point the domain structure is very interesting, exhibiting a power law, fractal, distribution of domain sizes. If we cool the system from a high temperature domain state, the MC procedure finds it very difficult to restore the magnetized state. This is also true in real magnets where a demagnetized magnet needs to be cooled in a field to restore its magnetic alignment. This is an example of non-ergodicity.

The mathematical singularity in the magnetization behavior only occurs for large samples at equilibrium, so that smaller samples have a rounded behavior near T_c. This rounding behaves as,

$$m(T_c, L) \approx L^{-\beta/\nu} \tag{78}$$

which is called finite size scaling. Moreover, the larger a sample is, the longer it takes to relax to its equilibrium value, and the relaxation time diverges near the critical point, a phenomenon called critical slowing down. So if the MC simulations runs are not long enough, the critical behavior is further rounded. This behavior is further exacerbated by non-ergodicity in the low temperature phase.

B. Classical molecular dynamics

God does not care about our mathematical difficulties. He integrates empirically. – Albert Einstein

Molecular dynamics (MD) covers a broad class of methods in both classical and quantum physics. Classical MD involves solving Newton's equations for each particle in the simulation box, while quantum or ab-intio MD involves solving for the time dependence of atoms or nuclei by first solving Schrodinger's equation for the relevant wavefunctions. We shall concentrate on classical MD. This seems trivial in that we simply solve Newton's equations to find the trajectory for each particle or atom in the system. The key difficulty is that we are usually interested in many particles, for example in simulating the evolution of Galaxies we need to take millions of stars, or in simulating a protein we need to take hundreds of different amino acid residues each of which has many atoms. The modern era of MD began with the work of Alder and Wainright on hard- sphere liquids in the late 1950s. The paper by Rahman in 1964 on a molecular dynamics simulation of liquid argon with a soft sphere (LennardJones) potential represented an important next step. Simulations of complex fluids followed; the original study of liquid water by Stillinger and Rahman was published in 1971. Simulations of proteins began soon after with to of the early innovators being Levitt and Karplus.

Now MD is a standard tool in all areas of science and engineering, and there are many MD packages available for a variety of different communities. For proteins and other biomolecules the standard packages are CHARM, AMBER, GROMOS... while for polymers and other materials LAMMPS is a standard. A key element in all of these packages is the interatomic potential that is used. The force is found from the potential and the range of the potential determines the type of algorithm that is used, with different methods being applied to long range as opposed to short range potentials. Many of these packages also have Monte Carlo (MC) options and some also have options for quantum calculations, i.e. quantum MC and/or quantum MD.

MD methods can be considered as a method for sampling the classical phase space density $\rho(\{q_i, p_i\}, t)$. The phase space density obeys the Liouville equation that could itself be integrated, but its phase space is very large so this approach is limited to a few particles. However, if ergodicity holds we carry out an MD calculation with any starting configuration and at long times all configurations will be sampled. In practice, ergodicity is difficult to achieve as the energy landscape of most interesting problems is rough, and because the fundamental timestep of the MD method for atomic systems is very short, typically of order a femtosecond. Even after millions of timesteps the total MD simulation time in atomic systems is only nanoseconds or at most microseconds, which is a fundamental limitation of the method. The basic MD algorithm is as follows.

```
Basic molecular dynamics algorithm
Choose the number of particles to be simulated $N$, the interaction potential and the boundary conditions.
Place the particles at initial positions $\vec{r}_i(t = 0)$.
Choose initial velocities for all of the particles $\vec{v}_i(t = 0)$.
Then, do
  1. Find the force on each particle $\vec{f}_i(t)$
  2. Use this force and the velocities $\vec{v}_i(t)$ to move each particle to a new position $\vec{r}_i(t + \delta t)$
  3. Calculate $\vec{f}_i(t + \delta t)$ using these new positions
  4. Find the new velocities, $\vec{v}_i(t + \delta t)$
  5. Calculate averages. If converged exit, otherwise return to 1
```

To carry out steps 2 and 4, we need a procedure to go from the acceleration (or force) to the positions and velocities. We also need to define a potential. First we look at the way that we find the positions and velocities from the force.

1. The Verlet algorithm

The molecular dynamics (MD) trajectories (or configurations) of classical particles are found by numerically integrating Newton's second law for each particle, ie.

$$\vec{F}_i = m_i \vec{a}_i = m_i \frac{d^2\vec{r}_i}{dt^2} \tag{79}$$

where in the discussion below we take all of the masses to be the same.

There are many ways to numerically integrate these equations and their accuracy is dependent on some power of the time interval used in the discrete approximation. The first derivative such as velocity may be written in three basic ways: the backward difference;

$$\vec{v}(t) = \frac{d\vec{r}}{dt} = \frac{\vec{r}(t) - \vec{r}(t-\delta t)}{\delta t} + O(\delta t) \tag{80}$$

the forward difference;

$$\vec{v}(t) = \frac{d\vec{r}}{dt} = \frac{\vec{r}(t+\delta t) - \vec{r}(t)}{\delta t} + O(\delta t) \tag{81}$$

and the central difference

$$\vec{v}(t) = \frac{d\vec{r}}{dt} = \frac{\vec{r}(t+\delta t) - \vec{r}(t-\delta t)}{2\delta t} + O(\delta t)^2 \tag{82}$$

Notice that the central difference has a correction of higher order than the forward or backward difference, as is seen by expanding the right hand side of these equations in a Taylor series. The second derivative can be written in a similar manner, however the best form is the central difference,

$$\vec{a}(t) = \frac{d^2\vec{r}(t)}{dt^2} = \frac{\vec{r}(t+\delta t) - 2\vec{r}(t) + \vec{r}(t-\delta t)}{\delta t^2} + O(\delta t)^2 \tag{83}$$

To demonstrate these equations we write the Taylor series,

$$\vec{r}(t+\delta t) = \vec{r}(t) + \delta t \vec{r}'(t) + \frac{1}{2!}\delta t^2 \vec{r}''(t) + \frac{1}{3!}\delta t^3 \vec{r}'''(t) + \frac{1}{4!}\delta t^4 \vec{r}''''(t) + .. \tag{84}$$

and

$$\vec{r}(t-\delta t) = \vec{r}(t) - \delta t \vec{r}'(t) + \frac{1}{2!}\delta t^2 \vec{r}''(t) - \frac{1}{3!}\delta t^3 \vec{r}'''(t) + \frac{1}{4!}\delta t^4 \vec{r}''''(t) + .. \tag{85}$$

For example the discrete or central difference form of the second derivative is found by adding these two equations. Using the central difference form of the second derivative in Newton's equations, we write the most basic form of the Verlet algorithm,

$$\vec{r}(t+\delta t) = 2\vec{r}(t) - \vec{r}(t-\delta t) + \delta t^2 \vec{F}_i(t)/m_i \quad \text{Basic Verlet} \tag{86}$$

In order to use this equation we need the positions as two times $\vec{r}(t)$ and $\vec{r}(t-\delta t)$. Alternatively we can use the backward difference for the velocity and write expressions in terms of the velocity at time t,

$$\vec{r}_i(t+\delta t) = \vec{r}_i(t) + \vec{v}_i(t)\delta t + \delta t^2 \vec{F}_i(t)/m_i \quad \text{Basic Verlet} \tag{87}$$

and

$$\vec{v}_i(t+\delta t) = \frac{\vec{r}_i(t+\delta t) - \vec{r}_i(t)}{\delta t} \tag{88}$$

This algorithm works quite well, but a better algorithm is found by using a so-called leapfrog procedure, which leads to the velocity Verlet algorithm, which is given by the two equations,

$$\vec{r}_i(t+\delta t) = \vec{r}_i(t) + \delta t \vec{v}_i(t) + \frac{1}{2}\delta t^2 \vec{F}_i(t)/m_i \tag{89}$$

and

$$\vec{v}_i(t+\delta t) = \vec{v}_i(t) + \frac{1}{2}\delta t(\vec{F}_i(t+\delta t) + \vec{F}_i(t)) \tag{90}$$

So step 2 in the MD algorithm above is given by Eq. () or (), while step 4 is given by Eq. () of (). Note that only one force calculation is required for each timestep, as the forces at the previous time can be stored, so only step 3 needs to be carried out at each timestep, while step 1 uses forces calculated at the previous timestep.

2. An example - MD model for Argon

The next step is to define a potential energy to describe the interactions between the particles. As an example, we consider an important interatomic potential that is used in materials science, chemistry and biology, the Lennard-Jones pair potential,

$$V(r_{ij}) = 4\epsilon[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}]. \tag{91}$$

The repulsive term is due to coulomb and exchange interactions of the electron clouds of each particle, while the attractive term is due to induced dipole-dipole interactions. For Argon the parameters of the model are $m = 6.69 \times 10^{-26} kg$, $\epsilon = 1.654 \times 10^{-21} J$, $\sigma = 3.405 \times 10^{-10} m$. The minimum of the potential is at $r_m = 2^{1/6}\sigma$, and the depth of the minimum is $-\epsilon$. The time scale of the problem is $\tau = (m\sigma^2/\epsilon) = 2.17ps$. In most simulations, scaled units are used, which in our problem corresponds to setting, $m = 1, \sigma = 1, \epsilon = 1$, so that one unit of time in our simulations corresponds to $2.17ps$. In practice, the timestep required for the simulations in the scaled units is $\delta t = .004$, which corresponds to a real time unit of $8.68fs$. This is typical of simulations of atomic motion and in cases where the interactions are strong, the timestep has to be shorter, for example in covalently bonded systems. Since the timestep is so short, simulations over millions of timesteps only lead to real times of order ns to μs, which is a fundamental limitation of MD simulations. The force between two atoms found from the LJ potential is,

$$\vec{F}(r_{ij}) = -\frac{\partial V}{\partial r_{ij}} \hat{r}_{ij} \tag{92}$$

The total force on a particle is found by summing this force over all neighbors of the particle.

With these procedures, we are able to follow the trajectories of all the particles in the simulation. From these trajectories, we would like to calculate properties. But what can we calculate. Since the calculation conserves the total energy, the temperature does not change (on average), so how do we change the temperature? How do we calculate the pressure and the equation of state? To understand this better we need to understand classical interacting particle systems more fully.

C. Finding T and P using energy conserving MD

In classical particle systems the Hamiltonian or Energy can often be approximated by pair interactions, so that

$$H = KE + V = \sum_{i=1}^{N} \frac{\vec{p}_i^2}{2m} + \sum_{i>j}^{N} V_2(\vec{r}_{ij}). \tag{93}$$

In the cannonical ensemble, the probability finding a particle at a point in phase space is given by,

$$p = Ae^{-\beta H} = p_r p_v. \tag{94}$$

where H is the Hamiltonian. Note that the Boltzmann probabilitiy factors into a part that contains the potential and a part that contains the momentum. The momentum part of the Hamilonian is simply $\sum_i \vec{p}_i^2/2m$. Now note that p_v also factors, i.e. $p_v = p_{v_1}...p_{v_N}$ and the probability of finding any one particle in a momentum state $\vec{p}_i$ is then,

$$p_{MB}(\vec{p}_i) = A_1 e^{-\beta\frac{\vec{p}_i^2}{2m}} = \left(\frac{\beta}{2m\pi}\right)^{3/2} e^{-\frac{\vec{p}_i^2}{2mk_BT}}. \tag{95}$$

This expression holds for every particle in the system. Often this is written in terms of the velocities

$$p_{MB}(\vec{v}) = \left(\frac{m\beta}{2\pi}\right)^{3/2} e^{-\frac{m\vec{v}^2}{2k_BT}} \tag{96}$$

where the constant A_1 is found by normalizing the probability. From this "Maxwell-Boltzmann" distribution it is easy to show that the root mean square velocity of particles is related to the temperature through

$$\frac{1}{2}m < \vec{v}^2 >= \frac{3}{2}k_BT \quad \text{or} \quad < KE >= \frac{3}{2}Nk_BT \qquad \text{for a monatomic system} \tag{97}$$

More generally, the kinetic energy per degree of freedom is $\frac{1}{2}k_BT$, which is of course the equipartition theorem. For an ideal gas the kinetic energy and the internal energy are the same. However for an interacting system the internal energy is the kinetic plus potential energy, so the equipartition result applies to the kinetic energy for systems with interactions that are momentum independent. Now note that from this result we have a procedure for changing the temperature in our MD simulations. We simply rescale the velocities to satisfy Eq. (). This is one of the most common "thermostats" used in MD simulations.

Now we want to calculate the equation of state for the interacting classical gas, i.e. we want to see how the ideal gas law is modified by the interactions. To do this, we use the virial method following standard definitions in mechanics, where we define the "virial" $G = \sum_k \vec{p}_k \cdot \vec{r}_k$, so that

$$\frac{dG}{dt} = 2K + \sum_{k=1}^{N} \vec{F}_k \cdot \vec{r}_k = 2K + W_{int} + W_{ext} \tag{98}$$

W_{int} is the work of the internal forces, while W_{ext} is the work of the external forces. The virial theorem states that if the average of the time derivative of G is zero, then,

$$< \frac{dG}{dt} >= 0 =< \sum_k \frac{\vec{p}_k^2}{m} > + \sum_{k=1}^{N} < \vec{F}_k \cdot \vec{r}_k >= 2 < KE > + \sum_{k=1}^{N} < \vec{F}_k \cdot \vec{r}_k > \tag{99}$$

The generalized theorem states that this equality is correct for each component of the sum above, as we shall show later. The reason that the average of the time derivative of G is zero at long time is that it is approximately $(G_2 - G_1)/(t_2 - t_1)$, with G_2 and G_1 bounded and with zero average.

$$\sum_{k=1}^{N} < \vec{F}_k \cdot \vec{r}_k >= \sum_{k=1,N}^{internal} < \vec{F}_k \cdot \vec{r}_k > + L_x(-PL_yL_z) + L_y(-PL_xL_z) + L_z(-PL_xL_y) \tag{100}$$

where P is the external pressure applied to the cubic box. We then have,

$$PV = Nk_BT + \frac{1}{3} \sum_{k=1,N}^{internal} < \vec{F}_k \cdot \vec{r}_k > \rightarrow Nk_BT + \frac{1}{3} \sum_{j>i} \vec{r}_{ij} \cdot \vec{F}_{ij} \tag{101}$$

where the last expression on the right is for the case of pair potentials, such as the LJ model for inert gases such as Argon.

From the MD simulation we get a series of configurations. From these configurations we can find the temperature and the pressure. We can also look at other more detailed properties such as the mean square displacement of particles, and the local arrangements of the particles. Since MD uses a realistic dynamics, the dynamical properties can also be studied, for example if we start with a dense cluster of Argon at high temperatures how long does it take to melt or evaporate? We can also do the opposite simulations, i.e. if we state with a gas of Argon and quench the system to low temperatures, does it form crystals of solid Argon? Does it form a glass?

There is an interesting and useful result used in many MD simulations, called the Jarzynski equality (JE), which is an equation that relates free energy differences between two equilibrium states and non-equilibrium processes. We shall return to this issue later in the course.

VI. THERMODYNAMIC POTENTIALS

A. Reversibility and exact differentials

If the differential of a scalar function of n variables $F(x_1...x_n)$ is given by,

$$dF = \sum_i f_i(x_1...x_n)dx_i \tag{102}$$

and a Taylor expansion exists so that,

$$f_i = \frac{\partial F}{\partial x_i}, \tag{103}$$

Then dF is an exact differential and,

$$\frac{\partial f_i}{\partial x_l} = \frac{\partial^2 F}{\partial x_i \partial x_l} = \frac{\partial f_l}{\partial x_i} = \frac{\partial^2 F}{\partial x_l \partial x_i} \tag{104}$$

That is, the order of the derivatives do not matter. These obvious looking identities are used extensively in thermodynamics and in that context are called Maxwell relations.

From a physics perspective, the most important property of an exact differential, is that its integral depends only on its end points, i.e. if F is an exact differential and i, f are arbitrary initial and final points of a path integral, then

$$\int_A^B dF = F_B - F_A \quad \text{or equivalently} \quad \oint dF = 0. \tag{105}$$

Clausius noted that at equilibrium dQ/T is an exact differential with these properties. In physics variables with these properties are called state variables. At equilibrium, all thermodynamic variables are state variables. Moreover, the second law states that if a processes is carried out away from equilibrium $dQ > TdS$, so that more heat is required to move the system from one state to another, hence there is a loss of efficiency in the use of the heat.

Since thermodynamic variables are exact differentials, a Taylor expansion exists, so that for example the internal energy $U(S, V, N)$ may be written as,

$$dU = TdS - PdV + \mu dN = \left(\frac{\partial U}{\partial S}\right)_{V,N} dS + \left(\frac{\partial U}{\partial V}\right)_{S,N} dV + \left(\frac{\partial U}{\partial N}\right)_{S,V} dN \tag{106}$$

so that,

$$T = \left(\frac{\partial U}{\partial S}\right); \quad -P = \left(\frac{\partial U}{\partial V}\right)_{S,N}; \quad \mu = \left(\frac{\partial U}{\partial N}\right)_{S,V} \tag{107}$$

Similar relations may be written down for any thermodynamic variable. Moreover it is straightforward to write down Maxwell relations from these expressions. Before doing that we need to understand more fully the free energies that are used in thermodynamics and their physical origins.

In a similar manner the entropy, $S(U, V, N)$, can be written as,

$$dS = \frac{1}{T}dU + \frac{P}{T}dV - \frac{\mu}{T}dN = \left(\frac{\partial S}{\partial U}\right)_{V,N} dU + \left(\frac{\partial S}{\partial V}\right)_{U,N} dV + \left(\frac{\partial S}{\partial N}\right)_{U,V} dN \tag{108}$$

so that,

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_{V,N}; \quad P = \left(\frac{\partial S}{\partial V}\right)_{U,N}; \quad -\frac{\mu}{T} = \left(\frac{\partial S}{\partial N}\right)_{U,V} \tag{109}$$

The volume, $V(U, S, N)$ may be written as,

$$dV = -\frac{1}{P}dU + \frac{T}{P}dS + \frac{\mu}{P}dN = \left(\frac{\partial V}{\partial U}\right)_{S,N} dU + \left(\frac{\partial V}{\partial S}\right)_{U,N} dS + \left(\frac{\partial V}{\partial N}\right)_{S,V} dN \tag{110}$$

so that,

$$-\frac{1}{P} = \left(\frac{\partial V}{\partial U}\right)_{S,N}; \quad \frac{T}{P} = \left(\frac{\partial V}{\partial S}\right)_{U,N}; \quad \frac{\mu}{P} = \left(\frac{\partial V}{\partial N}\right)_{U,S} \tag{111}$$

Finally, the number of particles, $N(U, S, V)$, has the representation,

$$dN = \frac{1}{\mu}dU - \frac{T}{\mu}dS + \frac{P}{\mu}dV = \left(\frac{\partial N}{\partial U}\right)_{S,V} dU + \left(\frac{\partial N}{\partial S}\right)_{U,V} dS + \left(\frac{\partial N}{\partial V}\right)_{U,S} dN \tag{112}$$

so that,

$$\frac{T}{\mu} = \left(\frac{\partial N}{\partial U}\right)_{S,V}; \quad -\frac{T}{\mu} = \left(\frac{\partial N}{\partial S}\right)_{U,V}; \quad \frac{P}{\mu} = \left(\frac{\partial N}{\partial V}\right)_{U,S} \tag{113}$$

In writing these expressions, the independent variables are one of the set U, S, V, N. Sometimes we need to use different independent variables, due to the way in which experiments or simulations are carried out. The set of variables U, S, V, N are most natural for the microcanonical ensemble. However in the canonical ensemble we want the independent variable T instead of S. In the grand canonical ensemble, in addition, we want the independent variable μ instead of N. One extension that we consider below is where we have a charged species. In that case the conjugate variables are charge Q and voltage V_0. In the microcanonical ensemble Q is the independent variable, while in the grand canonical ensemble V_0 is the independent variable. Actually the electrostatic potential can be incorporated as a contribution to the chemical potential, though we shall treat them separately for now.

B. Work done by the reservoir, Legendre transforms

One of the fundamental postulates of equilibrium statistical physics is that entropy is maximized. However we know that in mechanics and in many other ground state problems energy is minimized. Moreover if a system is connected to a reservoir, the work done by the reservoir also has to be taken into account. A simple example from electrostatics is to consider a capacitor that is initially uncharged attached to a battery. The battery is the reservoir in this case. Now close a switch to the battery. The lowest energy state of the capacitor is the uncharged state, however the battery supplies charge to the capacitor, reducing the energy in the battery. If the capacitor has capacitance C, and the battery has voltage V_0, the energy of the system at any charge state of the capacitor is,

$$E^{(t)} = \frac{Q^2}{2C} - QV_0 \tag{114}$$

Note that here we do not worry about the total energy in the reservoir, but only the change in energy of the reservoir ($-QV_0$) as the capacitor is charged. The first term is the energy in the capacitor and the second term is the energy supplied by the reservoir (the battery). Minimizing with respect to the charge gives $Q = CV_0$. In doing this minimization, we consider that the battery is perfect so that the voltage does not change. We also assume that there is no dissipation. In this example, the charge is "conjugate" to the voltage of the battery. The point of this example is that if we minimized the energy of the capacitor without the effect of the reservoir, the answer would be $Q = 0$, but in the presence of the reservoir, we get a different answer, $Q = CV_0$.

Now consider the mathematical definition of the Legendre transform. If we have a function of $f(\{x_i\})$, the Legendre transform with respect to the variable x_1 is,

$$L[f, x_1] = f - \frac{\partial f}{\partial x_1} x_1 \tag{115}$$

Clearly this can be repeated to generate a Legendre transform with respect to any number of the variables $\{x_i\}$. For a function of one variable, we write,

$$L(f(x)) = f - \frac{\partial f}{\partial x} x = f - f_x x \tag{116}$$

This can also be considered as the equation for the tangent to a line, where f is the intercept and f_x is the slope. The Legendre transform can then be considered to be a family of tangents to the original function. Note that $L(L(f)) = f$. To show this, we define $g = f - f_x x$, so that

$$L(g(f_x)) = g - g_{f_x} f_x = f - f_x x - \frac{\partial(f - f_x x)}{\partial(f_x)} f_x = f \tag{117}$$

There are many other intriguing properties of the Legendre transform, however it is better to keep in mind the physical reasons for introducing the Legendre transform.

C. The right free energy to minimize

In a statistical physics or thermodynamics calculations we can either maximize the entropy or minimize the correct free energy. In both cases we have to take into account the contribution of the reservoir. Many reservoirs can be considered. For example consider the internal energy that includes many terms in the generalized work,

$$dU = TdS - PdV + \vec{H} \cdot d\vec{M} + \vec{E} \cdot d\vec{P} + \sum_i \mu_i dN_i + V_i dQ_i + fdL + ... \tag{118}$$

where many physical processes are considered. In addition to the terms in Eq. (), a term for the case of charge transfer of multiple species ($V_i dQ_i$) and a simple elastic process (fdL) are added. The chemical potential and charge transfer pieces can be combined in many cases, but it is pedagogically nice to treat them separately for the time being. Now note that the quantities that are the independent variables in this expression, $S, V, \vec{M}, \vec{P}, N_i, Q_i, L$ are, in most experiments, NOT the things that are controlled. Instead the conjugate variables $P, \vec{H}, \vec{E}, \mu_i, V_i, f$ are most often the quantities varied. In statistical physics terms there is a reservoir that has this conjugate variable fixed. In these cases, the quantity that is minimized is the Legendre transform of the internal energy with respect to the quantities that are actually varied in the experiment. For example in the case of a charge reservoir for one species, we have $U \rightarrow U - V_i Q_i$. This can be compared to the example of a capacitor (). In general we can define an energy $Y(T, P, \mu, \vec{H}...)$ where the variables in brackets define the Legendre transform that is taken. For example,

$$Y(S, P, \vec{H}, N_i, f) = U + PV - \vec{H} \cdot \vec{M} - \sum_i \mu_i dN_i - fdL \tag{119}$$

Some of these Legendre transform functions are sufficiently important that they have been given their own names. We have already seen two of them. Consider the case of a canonical ensemble where the reservoir is at constant temperature. The heat energy supplied by the reservoir is then $-TS$. This is the physical origin of the definition of the Legendre transform $U - TS$ as the correct quantity that should be minimized. This is the Helmholtz free energy (A or F), so that

$$A = U - TS; \quad \text{so that} \quad dA = dU - TdS - SdT = -SdT - PdV + \sum_i \mu_i dN_i \tag{120}$$

where we used the fundamental theorem of thermodynamics to deduce the final expression on the right hand side. We have also found that $A = F = -k_B Tln(Z)$, where Z is the canonical partition function. Of course many experiments are carried out by controlling T rather than S, so the the Helmholtz free energy should be used in the theoretical study of optimal states. In some experiments, the control variable is the heat added to the system, so that the entropy is effectively the control variable, however the pressure is controlled instead of the volume as occurs in many gas phase processes with a piston. In the case where the entropy is controlled and there is the the possibility of volume change at fixed pressure, we define the enthalpy,

$$H = U + PV; \quad \text{so that} \quad dH = TdS + VdP + \sum_i \mu_i dN_i. \tag{121}$$

In this ensemble no heat is exchanged, but particles are exchanged. In cases where the reservoir allows heat exchange at constant T and volume change at constant P, we must minimize the Gibb's free energy which is the correct Legendre transform for that case, i.e.,

$$G = U + PV - TS = A + PV \quad \text{so that} \quad dG = -SdT + VdP + \sum_i \mu_i dN_i \tag{122}$$

From the relation $U = TS - PV + \mu N$, we find that $G = \mu N$, so the Gibb's free energy is the chemical potential per particle.

Finally, we have the grand potential which is the correct Legendre transform for a reservoir that allows exchange of heat and particles, but the volume is fixed, so the correct Legendre transform is,

$$\Phi_G = U - TS - \sum_i \mu_i N_i \quad \text{so that} \quad d\Phi_G = -SdT - PdV - \sum_i N_i d\mu_i, \tag{123}$$

If we use this relation with the equation $U = TS - PV + \mu N$, we find that $\Phi_G = -PV$. Recall that $\Phi_G = -k_B Tln(\Xi)$, so that $PV = k_B Tln(\Xi)$, so we can find the pressure and hence the equation of state from the grand canonical partition function. We will use this relation frequently when working with Bose and Fermi gases.

Note that we could carry through a similar procedure for the entropy, but it is unfamiliar as it is seldom used. The entropy in the microcanonical ensemble is changed when a connection is made to a reservoir, and the change is of a similar form to that required for add the effect of a reservoir to the internal energy.

Of course the entropies are always maximized while the energies (thermodynamic potentials) are minimized.

VII. RESPONSE FUNCTIONS, ENERGY AND PARTICLE NUMBER FLUCTUATIONS

A. Thermodynamic response functions

Thermodynamic response functions are often measured as they provide excellent indicators of phase changes and access to fundamental constants. Moreover, as we shall see below, response functions are related to fluctuations of the energy and particle number. Response functions are not all independent and there are relations between them. These relations provide important checks on theoretical and experimental results.

Five commonly measured thermodynamic response functions are: the specific heat at constant volume C_V or at constant pressure C_P, the adiabatic and isothermal compressibilities, κ_S, κ_T and the coefficient of thermal expansion α_P. These quantities are defined through the relations,

$$C_V = \left(\frac{\partial Q}{\partial T}\right)_{V,N} = \left(\frac{\partial U}{\partial S}\right)_{V,N}\left(\frac{\partial S}{\partial T}\right)_{V,N} = T\left(\frac{\partial S}{\partial T}\right)_{V,N}, \tag{124}$$

$$C_P = \left(\frac{\partial H}{\partial T}\right)_{P,N} = \left(\frac{\partial H}{\partial S}\right)_{P,N}\left(\frac{\partial S}{\partial T}\right)_{P,N} = T\left(\frac{\partial S}{\partial T}\right)_{P,N}, \tag{125}$$

$$\kappa_S = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_{S,N} = -\left(\frac{\partial lnV}{\partial P}\right)_{S,N}, \tag{126}$$

$$\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_{T,N} = -\left(\frac{\partial lnV}{\partial P}\right)_{T,N}, \tag{127}$$

$$\alpha_P = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_{P,N} = \left(\frac{\partial lnV}{\partial T}\right)_{P,N}, \tag{128}$$

B. Energy and particle number fluctuations, central limit theory and relation to response functions

In the microcanonical ensemble, the energy is fixed, however there are fluctuations in the kinetic energy and in the potential energy. Since the kinetic energy fluctuates, the temperature fluctuates. In the canonical ensemble, the temperature is fixed but the energy fluctuates. In the grand-canonical ensemble, the energy and the number of particles both fluctuate. In this section we will explore the relations of fluctuations to the thermodynamic variables in the problem. We shall focus on fluctuations in the canonical and grand canonical ensemble.

First consider the canonical ensemble where the temperature is fixed but the energy fluctuates. The internal energy is the mean of the energy, while the fluctuations in the energy are related to the size of the system as demonstrated by the central limit theorem, from which we know $\delta E/E \sim 1/N^{1/2}$. However it is possible to find a much more precise relation by studying the energy fluctuations directly. To do this we first define,

$$\delta E = [<(E-<E>)^2>]^{1/2} = [<E^2> - <E>^2]^{1/2} \tag{129}$$

where we use the notation that $<O>=\sum_i O_i p_i$. Now note that,

$$<E>= U = \frac{-1}{Z}\frac{\partial Z}{\partial \beta}; \quad \text{and} \quad <E^2>= \frac{1}{Z}\frac{\partial^2 Z}{\partial \beta^2} \tag{130}$$

From these expressions we find the neat result,

$$(\delta E)^2 =<E^2> - <E>^2 = \frac{1}{Z}\frac{\partial^2 Z}{\partial \beta^2} - (\frac{1}{Z}\frac{\partial Z}{\partial \beta})^2 = \frac{\partial^2 ln(Z)}{\partial \beta^2}. \tag{131}$$

Therefore we find that,

$$(\delta E)^2 = -\left(\frac{\partial U}{\partial \beta}\right)_{V,N} = -\left(\frac{\partial U}{\partial T}\right)_{V,N}\left(\frac{\partial T}{\partial \beta}\right)_{V,N} = k_B T^2 C_V \tag{132}$$

Note that it is easy to show that the specific heat may also be written as,

$$C_V = T\left(\frac{\partial S}{\partial T}\right)_{V,N} = -T\left(\frac{\partial^2 F}{\partial T^2}\right)_{V,N}. \tag{133}$$

The square of the energy fluctuations are thus proportional to the specific heat. This gives a nice way of getting the specific heat from a numerical simulation in the canonical ensemble. An alternative method is to calculate $U =< E >$ and then take a numerical derivative with respect to T, however this is not as accurate computationally.

Since the energy and the specific heat are extensive, we have,

$$\frac{\delta E}{E} \propto \frac{1}{N^{1/2}} \tag{134}$$

as required by the central limit theorem. This means that the energy calculated in the canonical ensemble can deviate at most by $N^{1/2}$ from the energy used in the microcanonical ensemble, so the thermoynamic behavior is the same.

A similar argument can be carried out for the fluctuations in the particle number within the grand canonical ensemble, so that,

$$(\delta N)^2 =< N^2 > - < N >^2 = (k_BT)^2\frac{\partial^2 ln(\Xi)}{\partial \mu^2} = k_BTV\left(\frac{\partial^2 P}{\partial \mu^2}\right)_{T,V} \tag{135}$$

after some analysis this may be related to the isothermal compressibility κ_T,

$$\frac{(\delta N)^2}{N} = k_BT\rho\kappa_T \tag{136}$$

where $\rho = N/V$ is the number density. Again the central limit theorem is evident as $(\delta N)^2/N \approx 1$, so $\delta N/N \approx 1/N^{1/2}$.

It is important to note that we can also find the specific heat in the grand-canonical ensemble, and the result will be the same as the cannonical ensemble, up to deviations of order $1/N^{1/2}$. We can even find the specific heat from the microcanonical ensemble where it is related to the fluctuations in the kinetic energy (the Lebowitz equation), which is often used in MD simulations.

C. Maxwell relations and finding expressions in terms of response functions

Response functions are the most commonly measured properties of materials. In the case of particle systems, commonly measured response functions include, C_V, C_P, κ_S, κ_T and α_P (see Eq. (124-128) for the definitions of these response functions. Note that all of these response functions are defined at constant N, so we are working with two variable expressions. In more general cases a formulation in terms of Jacobians can be useful, though there is some notation overhead to learn. In the case of magnetic systems the magnetic susceptibility, χ, is the most commonly measured response function. Here we shall discuss methods to find relations between response functions and to reduce most thermodynamic derivatives to expressions involving the standard response functions above.

Maxwell relations come from second derivatives of the free energies. Using the derivative relations for each of the standard free energyies and the entropy ie. S, U, F, H, G, it is easy to write down three Maxwell relations for each case, leading to 15 Maxwell relations in all. See Appendix H of Pathria and Beale. For example for the internal energy, we start with the fundamental thermodynamic relation to write,

$$dU = TdS - PdV + \mu N = \left(\frac{\partial U}{\partial S}\right)_{V,N} dS + \left(\frac{\partial U}{\partial V}\right)_{S,N} dV + \left(\frac{\partial U}{\partial N}\right)_{S,V} dN \tag{137}$$

so that,

$$T = \left(\frac{\partial U}{\partial S}\right); \quad -P = \left(\frac{\partial U}{\partial V}\right)_{S,N}; \quad \mu = \left(\frac{\partial U}{\partial N}\right)_{S,V} \tag{138}$$

and hence

$$\left(\frac{\partial^2 U}{\partial S \partial V}\right) = \left(\frac{\partial^2 U}{\partial V \partial S}\right) \quad \text{so} \quad \left(\frac{\partial T}{\partial V}\right)_{S,N} = -\left(\frac{\partial P}{\partial S}\right)_{V,N} \tag{139}$$

The procedure to derive any of the Maxwell relations is then: (i) Look at the independent variables in the expression. (ii) choose the free energy with these independent variables. (iii) write down the derivative form of this free energy and hence the relations between dependent variables and derivatives of the free energy, (iv) Choose the appropriate cross derivatives to reproduce the Maxwell relation.

For example let's prove the Maxwell relation,

$$\left(\frac{\partial \mu}{\partial P}\right)_{S,N} = \left(\frac{\partial V}{\partial N}\right)_{S,P} \tag{140}$$

The enthalpy $H(P,S,N) = U + PV$ has the appropriate independent variables. We have,

$$dH = dU + PdV + VdP = TdS + VdP + \mu dN = \left(\frac{\partial H}{\partial S}\right)_{P,N} dS + \left(\frac{\partial H}{\partial P}\right)_{S,N} dP + \left(\frac{\partial H}{\partial N}\right)_{S,P} dN \tag{141}$$

so that,

$$\mu = \left(\frac{\partial H}{\partial N}\right)_{S,P} \quad \text{and} \quad V = \left(\frac{\partial H}{\partial P}\right)_{S,N} \tag{142}$$

and therefore,

$$\left(\frac{\partial \mu}{\partial P}\right)_{S,N} = \left(\frac{\partial^2 H}{\partial P \partial N}\right) \tag{143}$$

and

$$\left(\frac{\partial V}{\partial N}\right)_{S,P} = \left(\frac{\partial^2 H}{\partial N \partial P}\right). \tag{144}$$

The order of the partial derivatives on the RHS of (143) and (144) can be switched, therefore the Maxwell equation (140) is correct.

D. Solving strategies

Sometimes the independent variables in a thermodynamic expression are not the same as those in the standard potentials. In that case it is useful to define a new potential with the correct independent variables. For exampmle, consider Assigned Problem (14) which asks that we prove the relation,

$$T\left(\frac{\partial N}{\partial T}\right)_{V,\beta\mu} = \left(\frac{\partial E}{\partial \mu}\right)_{T,V}. \tag{145}$$

where $\beta = 1/T$ for this problem. The fact that $\beta\mu$ is held constant is a hint that an entropy form might help. However the natural variables for entropy are $S(U,V,N)$, and

$$dS = \frac{1}{T}dU + \frac{P}{T}dV - \frac{\mu}{T}dN \tag{146}$$

We want an entropy expression where $T, V, \beta\mu$ are held constant. However we can also use $\beta, V, \beta\mu$ and then use the chain rule for the T derivatives. To get an entropy expression with these independent variables, we write down the Legendre transformed entropy,

$$Y(\beta, V, \beta\mu) = S - \beta U + \beta\mu N \tag{147}$$

Writing an expression for dY and finding the appropriate Maxwell relation for the partials of Y, along with use of the chain and inverse rule (see below), solves the problem.

E. Partial derivative relations that help in reducing thermodynamic relations

The following rules of multivariable calculus can be used to help solve many problems in thermodynamics,

$$\left(\frac{\partial x}{\partial y}\right)_z = 1/\left(\frac{\partial y}{\partial x}\right)_z \qquad \text{Inversion} \tag{148}$$

$$\left(\frac{\partial x}{\partial y}\right)_z = \left(\frac{\partial x}{\partial w}\right)_z\left(\frac{\partial w}{\partial y}\right)_z \qquad \text{Addition of a variable} \tag{149}$$

$$\left(\frac{\partial x}{\partial y}\right)_z\left(\frac{\partial y}{\partial z}\right)_x\left(\frac{\partial z}{\partial x}\right)_y = -1 \quad \text{Triple Product} \tag{150}$$

If $A(x, y)$, then,

$$\left(\frac{\partial A}{\partial x}\right)_z = \left(\frac{\partial A}{\partial x}\right)_y + \left(\frac{\partial A}{\partial y}\right)_x\left(\frac{\partial y}{\partial x}\right)_z + \qquad \text{non} - \text{natural derivative} \tag{151}$$

Proofs:

1. Inversion: Consider $x(y, z)$ and $y(x, z)$, so that,

$$dx = \left(\frac{\partial x}{\partial y}\right)_z dy + \left(\frac{\partial x}{\partial z}\right)_y dz; \quad dy = \left(\frac{\partial y}{\partial x}\right)_z dx + \left(\frac{\partial y}{\partial z}\right)_x dz \tag{152}$$

In the case that $dz = 0$ we find from these two expressions,

$$dx = \left(\frac{\partial x}{\partial y}\right)_z dy; \quad dy = \left(\frac{\partial y}{\partial x}\right)_z dx \;\text{ so }\; \left(\frac{\partial x}{\partial y}\right)_z = 1/\left(\frac{\partial y}{\partial x}\right)_z \tag{153}$$

2. Addition of a variable:

$$\left(\frac{\partial x}{\partial y}\right)_z = \left(\frac{\partial x}{\partial w}\right)_z\left(\frac{\partial w}{\partial y}\right)_z \tag{154}$$

This is the inverse of the chain rule.

3. Triple product: If we want to find a triple product relation on three variables x, y, z that obey an equation of state. This is typical in thermodynamics, for example $PV = Nk_BT$ is really a function of three variables as we can define density $\rho = N/V$, so that $P = \rho k_BT$. We define a function $f(x, y, z) = 0$ whose solution leads an equation of state. If f has a Taylor expansions, we can write,

$$df = \left(\frac{\partial f}{\partial x}\right)_{yz} dx + \left(\frac{\partial f}{\partial y}\right)_{xz} dy + \left(\frac{\partial f}{\partial z}\right)_{xy} dz \tag{155}$$

Since $f(x, y, z) = 0$, $df = 0$. Moreover, this Taylor expansion must hold for any infinitesimal dx, dy, dz. We consider just the special cases $dx = 0$ or $dy = 0$ or $dz = 0$. In these cases we find,

$$\left(\frac{\partial x}{\partial y}\right)_z = -\frac{\left(\frac{\partial f}{\partial y}\right)_{xz}}{\left(\frac{\partial f}{\partial x}\right)_{yz}}; \quad \left(\frac{\partial z}{\partial x}\right)_y = -\frac{\left(\frac{\partial f}{\partial x}\right)_{yz}}{\left(\frac{\partial f}{\partial z}\right)_{xy}}; \quad \left(\frac{\partial y}{\partial z}\right)_x = -\frac{\left(\frac{\partial f}{\partial z}\right)_{xy}}{\left(\frac{\partial f}{\partial y}\right)_{xz}} \tag{156}$$

A product of these expressions leads to the triple product rule.

4. Non-natural derivative: Consider that we have a function $A(x, y)$ so that,

$$dA = \left(\frac{\partial A}{\partial x}\right)_y dx + \left(\frac{\partial A}{\partial y}\right)_x dy \tag{157}$$

Now divide through by dx while keeping a new variable z fixed, so that,

$$\left(\frac{\partial A}{\partial x}\right)_z = \left(\frac{\partial A}{\partial x}\right)_y + \left(\frac{\partial A}{\partial y}\right)_x\left(\frac{\partial y}{\partial x}\right)_z \tag{158}$$

F. Relating thermodynamic expressions to response functions

With some effort, most thermodynamic relations can be written in terms of measurable quantities, such as response functions. Besides the basic Maxwell relations outlined above, a variety of other derivatives may be stated, in particular in terms of "non-natural" variables. An example of an application of the non-natural derivative is as follows: The natural variables for the entropy are $S(U, V, N)$, but we may still consider the function $S(T, V)$, and assuming that N is constant we may write,

$$dS = \left(\frac{\partial S}{\partial T}\right)_V dT + \left(\frac{\partial S}{\partial V}\right)_T dV \tag{159}$$

Moreover, we may define new derivatives by dividing through by one of the infinitesimals (in this case dT or dV), and we may choose to keep a new variable constant, for example,

$$\left(\frac{\partial S}{\partial T}\right)_P = \frac{C_P}{T} = \frac{C_V}{T} + \left(\frac{\partial S}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_P = \frac{C_V}{T} + V\alpha_P \left(\frac{\partial S}{\partial V}\right)_T = \frac{C_V}{T} + V\alpha_P \left(\frac{\partial P}{\partial T}\right)_V \tag{160}$$

where a Maxwell relation

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial^2 F}{\partial T \partial V}\right) = \left(\frac{\partial P}{\partial T}\right)_V \tag{161}$$

was used to find the last relation on the RHS. Now we use the triple product,

$$\left(\frac{\partial P}{\partial T}\right)_{V,N} \left(\frac{\partial T}{\partial V}\right)_{P,N} \left(\frac{\partial V}{\partial P}\right)_{T,N} = -1 \tag{162}$$

so that,

$$\left(\frac{\partial P}{\partial T}\right)_{V,N} = -\frac{\left(\frac{\partial V}{\partial T}\right)_{P,N}}{\left(\frac{\partial V}{\partial P}\right)_{T,N}} = \frac{V\alpha_P}{V\kappa_T} \tag{163}$$

so that,

$$\left(\frac{\partial P}{\partial T}\right)_V = \frac{\alpha_P}{\kappa_T} \tag{164}$$

so we find the nice relation,

$$C_P = C_V + \frac{TV\alpha_P^2}{\kappa_T} \tag{165}$$

Another interesting relation between response functions is,

$$\kappa_S = \frac{C_V}{C_P}\kappa_T = \kappa_T - \frac{TV\alpha_P^2}{C_P}. \tag{166}$$

To show this, we use the triple product rule, then the inversion rule, then the chain rule, then inversion again and finally the triple product rule again. In a very succinct form this can be written as follows, (assuming N fixed),

$$\kappa_S = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_S = -\frac{1}{V}[\left(\frac{\partial S}{\partial V}\right)_P \left(\frac{\partial P}{\partial S}\right)_V]^{-1} = -\frac{1}{V}\frac{\left(\frac{\partial S}{\partial P}\right)_V}{\left(\frac{\partial S}{\partial V}\right)_P} = -\frac{1}{V}\frac{\left(\frac{\partial S}{\partial T}\right)_V \left(\frac{\partial T}{\partial P}\right)_V}{\left(\frac{\partial S}{\partial T}\right)_P \left(\frac{\partial T}{\partial V}\right)_P} = \frac{C_V}{C_P}\kappa_T \tag{167}$$

Using (165), this reduces to the last expression on the RHS of (166).

Assigned problems and sample quiz problems

Sample Quiz Problems

Quiz Problem 1. Prove the expression for the Carnot efficiency for a perfectly reversible Carnot cycle using an ideal gas.

Quiz Problem 2. Show that in the eigenfunction basis, the von Neumann entropy $S = -k_B tr(\hat{\rho} ln(\hat{\rho}))$ reduces to the Gibbs form $S = -k_B \sum_i p_i ln(p_i)$ where the sum is over all eigenstates.

Quiz Problem 3. Use Stirling's approximation to show that,

$$ln(\binom{N}{n}) \approx -N(pln(p) + (1-p)ln(1-p)), \quad \text{where} \quad p = n/N \tag{168}$$

Quiz Problem 4. If $S = -k_B N[pln(p) + (1-p)ln(1-p)]$, by doing a variation with respect to p find the value of p that gives the maximum entropy. Demonstrate that it is a maximum by showing that the second derivative with respect to p is negative.

Quiz Problem 5. Use the Master equation to prove the second law of thermodynamics, i.e. in a closed system $dS/dt \geq 0$.

Quiz Problem 6. Give three examples of systems where the ergodic hypothesis fails. Explain why it fails in these cases.

Quiz Problem 7. Why is the ideal gas law still a good starting point for high temperature gases, even though a non-interacting gas is non-ergodic?

Quiz Problem 8. Using $p_i = e^{-\beta E_i}/Z$ in the Gibbs form for the entropy, show that $F = -k_B T ln(Z)$, where $F = U - TS$ is the Helmholtz free energy. Here $Z = \sum_i e^{-\beta E_i}$ is the canonical partition function.

Quiz Problem 9. If $x = e^{-\beta(E_i - E_j)}$, so that $w_{ji} \to f(x)$. Show that the Metropolis algorithm is given by $f(x) = min(x, 1)$.

Quiz Problem 10. If $x = e^{-\beta(E_i - E_j)}$, so that $w_{ji} \to f(x)$. Show that detailed balance is satisfied for any $f(x)$ satisfying $f(x)/f(1/x) = x$. Moreover, show that $f(x) = x/(1+x)$, the heat bath algorithm, is one such solution.

Quiz Problem 11. For a monatomic interacting classical gas, with interactions that only depend on the particle co-ordinates, derive the Maxwell Boltzmann distribution of velocities and show that the average kinetic energy is given by $< KE >= 3Nk_BT/2$.

Quiz Problem 12. Using the fact that $\delta E^2 =< E^2 > - < E >^2 = k_B T^2 C_V$ show that $\delta E/E$ is proportional $1/N^{1/2}$.

Quiz Problem 13. Write down the central difference form of the second derivative. Using this expression for the acceleration, write down the basic form of the Verlet algorithm for Molecular dynamics.

Quiz Problem 14. Given that the virial is defined to be $G = \sum_k \vec{p}_k \cdot \vec{r}_k$. Explain why the average of its time derivative is expected to be zero at long times.

Quiz Problem 15. The ground state structure of Argon is a fcc crystal. If a MD simulation is carried out where a high temperature structure is quenched to a low temperature where the fcc structure is the equilibrium state. If the starting high temperature state in your MD simulation is a liquid phase, do you expect the simulation to find the correct fcc state? Explain.

Quiz Problem 16. In a MC calculation of the ferromagnetic Ising model on a square lattice, you calculate the magnetization as a function of temperature as accurately as you can. However your simulations remain rounded at the critical point, instead of showing the behavior $m \sim |T_c - T|^\beta$ expected in the thermodynamic limit. Explain the physical origins of the rounding that you observe. Your experimental colleague does a measurement of a model Ising magnet that consists of quantum dots with around 1000 spins in each quantum dot. She also observes a rounding of the behavior of the magnetization. Provide some reasons why her magnetization experiment also gives a rounded behavior near the critical point.

Problem 17. Derive the relation,

$$PV = Nk_BT + \frac{1}{3}\sum_{k}^{internal} \vec{r}_k \cdot \vec{F}_k \tag{169}$$

Explain the physical origin of each of the three terms. This relation is not always true for periodic boundary conditions, though rather fortuitiously it is true for periodic boundaries provided the interaction potential is a pair potential.

Problem 18. Show that

$$(\delta E)^2 = k_BT^2C_V \tag{170}$$

Assigned problems

Problem 1. Find the relation between pressure and volume for an ideal gas under going an adiabatic process.

Problem 2: Derive the ideal monatomic gas law using kinetic theory. You may use the equipartition result $U = \frac{3}{2}Nk_BT$.

Problem 3. Find the degeneracy $\Omega(E,N)$ for N Ising spins (spin 1/2) with Hamiltonian $H = -h\sum_i S_i$.

Problem 4. Consider a random walk on a one dimensional lattice where the probability of stepping to the right or to the left is 1/2. Find the probability that after N steps the walker is at position x lattice units from the starting position. Use Stirling's approximation to show that for large N the probability distribution is a Gaussian in x. Find the average mean square distance the walker is from the origin as a function of N.

Problem 5. By taking a time derivative of the definition of the density operator in terms of wavefunctions, $\hat{\rho}(t) = \sum_j |\psi_j(t) >< \psi_j(t)|$, derive the von Neumann time equation (12) of the lecture notes.

Problem 6. Prove the central limit theorem for a set of N independent subsystems that each have an energy E_i drawn from a Gaussian distribution with standard deviation σ.

Problem 7. Find the entropy of mixing for a system consisting of m different atom types, with n_i of each atom type and with the total of atoms, $N = \sum_i^m n_i$.

Problem 8. Derive Stirling's approximation $ln(N!) = Nln(N) - N$.

Problem 9. Which of the following is an exact differential? a) $dU(x,y) = 2xdx + 3xydy$; b) $dU(x,y) = (-1/y)dx + (x/y^2)dy$; c) $dU(x,y) = -ydx + xdy$? For these three examples, set $dU = 0$ and solve the differential equations to find $y(x)$. In cases where the differential is exact, find $U(x,y)$.

Problem 10. By using the fact that S is extensive show that

$$N\left(\frac{\partial S}{\partial N}\right)_{V,E} + V\left(\frac{\partial S}{\partial V}\right)_{N,E} + E\left(\frac{\partial S}{\partial E}\right)_{N,V} = S \tag{171}$$

and hence that $N\mu = E + PV - TS$.

Problem 11. Using the fact that $G = \mu N$ and $N\mu = E + PV - TS$ prove the Gibbs-Duhem relation.

Problem 12. From the fundamental thermodynamic relation show that,

$$\left(\frac{\partial \mu}{\partial P}\right)_{S,N} = \left(\frac{\partial V}{\partial N}\right)_{S,P} . \tag{172}$$

Problem 13. From the fundamental thermodynamic relation show that,

$$\left(\frac{\partial C_P}{\partial P}\right)_{T,N} = -T\left(\frac{\partial^2 V}{\partial T^2}\right)_{P,N} . \tag{173}$$

Problem 14. From the fundamental thermodynamic relation show that (here $\beta = 1/T$),

$$T\left(\frac{\partial N}{\partial T}\right)_{V,\beta\mu} = \left(\frac{\partial E}{\partial \mu}\right)_{T,V} \tag{174}$$

Problem 15. From the fundamental thermodynamic relation show that (here $\beta = 1/T$)

$$\frac{1}{T^2}\left(\frac{\partial T}{\partial \beta\mu}\right)_{E,V} = -\left(\frac{\partial N}{\partial E}\right)_{\beta\mu,V} \quad \text{and} \quad -\rho^2\left(\frac{\partial (S/N)}{\partial \rho}\right)_{T,N} = \left(\frac{\partial P}{\partial T}\right)_{\rho,N} \tag{175}$$

where $\rho = N/V$ is the number density.

Problem 16. From the results of problem 15 show that

$$\left(\frac{\partial P}{\partial \rho}\right)_{S/N,N} = \left(\frac{\partial P}{\partial \rho}\right)_{T,N} + [\left(\frac{\partial P}{\partial T}\right)_{\rho,N}]^2 \frac{NT}{\rho^2 C_V} \tag{176}$$

Problem 17. Consider a system with an order parameter, x. If the volume(V), energy(E) and particle number(N) are fixed, then the system will choose the value of x with the maximum entropy. Now consider that the same system is connected to a reservoir that has temperature T and chemical potential μ, so it can exchange energy and particles with the reservoir. Show that the total entropy, that is the sum of the entropy of the reservoir and the system, is maximized when

$$\left(\frac{\partial P}{\partial x}\right)_{T,\mu,V} = 0 \tag{177}$$

Hint: You can assume the system and reservoir are at the same temperature and chemical potential as required by equilibrium. Moreover you may use $PV = TS - E + \mu N$.

Problem 18. Consider a system where the thermodynamics are a function of temperature and another variable x which can be varied by an experimentalist. Assuming that

$$\left(\frac{\partial E}{\partial S}\right)_x = T; \quad \text{and using} \quad F = E - TS \tag{178}$$

show that,

$$\left(\frac{\partial E}{\partial x}\right)_S = \left(\frac{\partial F}{\partial x}\right)_T \tag{179}$$

Problem 19. Show that

$$C_V = -T\left(\frac{\partial^2 F}{\partial T^2}\right)_{N,V} \tag{180}$$

Using this relation show that, $< E^2 > - < E >^2 = k_B T^2 C_V$.

Problem 20. Show that

$$U = \left(\frac{\partial (F/T)}{\partial (1/T)}\right)_{N,V} \tag{181}$$

Problem 21. Show that

$$V\left(\frac{\partial P}{\partial T}\right)_{\mu,N} = S; \quad \text{and} \quad V\left(\frac{\partial P}{\partial \mu}\right)_{T,N} = N \tag{182}$$

Problem 22. From the fundamental thermodynamic relation show that

$$\left(\frac{\partial S}{\partial P}\right)_{T,\mu} = -\left(\frac{\partial V}{\partial T}\right)_{P,\mu} \tag{183}$$