Statistical Mechanics, Math 266: Week 1 Notes

January 5 and 7, 2010

1 The Microcanonical Ensemble

The Microcanonical Ensemble refers to a choice of a probability distribution on state space in which all configurations of a fixed energy E are given equal weight. To be concrete, we illustrate with a simple example. Consider a system of N classical point particles of mass m in a finite box $\Lambda = [0, L]^3 \subset \mathbb{R}^3$. The phase space of this system is given by $\Gamma = [0, L]^{3N} \times \mathbb{R}^{3N}$ where

$$(\mathbf{q}, \mathbf{p}) = (q_1, \ldots, q_N, p_1, \ldots, p_N) \in \Gamma$$

and $q_i \in [0, L]^3$ and $p_i \in \mathbb{R}^3$ for all $1 \leq i \leq N$. The energy of the system is given by the Hamiltonian, $H : \Gamma \to \mathbb{R}$ given by

$$H(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^{N} \frac{1}{2m} |p_i|^2 \tag{1}$$

The Hamiltonian enables us to define the submanifold of constant energy, Γ_E:

$$\Gamma_E = \{(\mathbf{q}, \mathbf{p}) \in \Gamma | H(\mathbf{q}, \mathbf{p}) = E\} \tag{2}$$

Hamiltonian dynamics is given by Hamilton's equations of motion:

$$\dot{q}_i = \frac{\partial}{\partial p_i} H(\mathbf{q}, \mathbf{p}) \tag{3}$$

$$\dot{p}_i = -\frac{\partial}{\partial q_i} H(\mathbf{q}, \mathbf{p}) \tag{4}$$

The following theorem of Liouville, stated without proof, will be necessary to define a measure on Γ_E.

Theorem 1.1. *Liouville*

The measure $d\boldsymbol{q}d\boldsymbol{p}$ on Γ is preserved under the Hamiltonian flow solving equations 3 and 4.

What this means precisely is the following. Consider the flow, $\varphi_t : \Gamma \to \Gamma$ such that $\varphi_t(\mathbf{q}_0, \mathbf{p}_0) = (\mathbf{q}(t), \mathbf{p}(t))$ is a solution of equations 3 and 4 when $\mathbf{q}(0) = \mathbf{q}_0$ and $\mathbf{p}(0) = \mathbf{p}_0$. Then letting B be a Lebesgue measurable set, define

$$B_t = \{\varphi_t(\mathbf{q}, \mathbf{p}) | (\mathbf{q}, \mathbf{p}) \in B\}$$

Liouville's theorem then says that

$$\mathrm{vol}(B_t) = \mathrm{vol}(B) \text{ for all } t \tag{5}$$

i.e.,

$$\int_{B_t} d\mathbf{q}d\mathbf{p} = \int_B d\mathbf{q}d\mathbf{p} \tag{6}$$

In our concrete example, this is easy to see, since

$$p_i(t) = p_i(0) \tag{7}$$

$$q_i(t) = q_i(0) + \frac{1}{m} p_i t \tag{8}$$

is just translation which clearly preserves volume.

Another important property is that H, energy, is conserved.

$$H(\mathbf{q}(t), \mathbf{p}(t)) = H(\mathbf{q}(0), \mathbf{p}(0)) \text{ for all } t \tag{9}$$

This is a direct consequence of the form of the Hamilton equations of motion:

$$\frac{d}{dt} H(\mathbf{q}(t), \mathbf{p}(t)) = \sum_{i=1}^{N} \frac{\partial H}{\partial q_i} \frac{dq_i}{dt} + \frac{\partial H}{\partial p_i} \frac{dp_i}{dt} \tag{10}$$

$$= \sum_{i=1}^{N} \frac{\partial H}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial H}{\partial p_i} \frac{\partial H}{\partial q_i} = 0 \tag{11}$$

This means that Γ_E, the manifold of constant energy E is invariant under the flow, i.e.,

$$\text{For all } (\mathbf{q}, \mathbf{p}) \in \Gamma_E \;,\; \varphi_t(\mathbf{q}, \mathbf{p}) \in \Gamma_E \tag{12}$$

It follows also that the trace measure of $d\mathbf{q}d\mathbf{p}$ on Γ_E will also be preserved under the flow φ_t. It may be intuitively clear what the "trace measure" means, but we provide a mathematical definition.

Let $B \subset \Gamma_E$ be a nice set and suppose Γ_E is a differentiable submanifold of Γ so that we define the normal to each point of Γ_E. Consider the normals at each point of B and define

$$B(\Delta E) = \left\{ (\mathbf{q}, \mathbf{p}) \in \Gamma \;\middle|\; \begin{array}{l} (\mathbf{q}, \mathbf{p}) \text{ lies on a normal to a point in } B \subset \Gamma_E \\ \text{and } H(\mathbf{q}, \mathbf{p}) \in [E, E + \Delta E] \end{array} \right\} \tag{13}$$

and

$$\nu_E(B) = \lim_{\Delta E \searrow 0} \frac{\mathrm{vol}(B(\Delta E))}{\Delta E} \tag{14}$$

It can be shown that this defines a measure of Γ_E which will clearly be invariant under the flow φ_t. In our case, Γ_E is compact, and

$$\nu_E(\Gamma_E) = \frac{d}{dE}\,\mathrm{vol}(\{(\mathbf{q},\mathbf{p})|H(\mathbf{q},\mathbf{p}) \leq E\}) \equiv \Omega_E \tag{15}$$

which we can easily compute:

$$\Gamma_E = [0,L]^{3N} \times \mathbb{S}^{3N-1}(\sqrt{2mE})\nu_E(\Gamma_E) = \frac{d}{dE}\left\{L^{3N}(2mE)^{\frac{3N}{2}}\omega_{3N}\right\} \tag{16}$$

where ω_d is the volume of the unit ball in $\mathbb{R}^d$ which is given by

$$\omega_d = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)} \tag{17}$$

So

$$\Omega_E = \nu_E(\Gamma_E) = 2mL^{3N}\pi^{\frac{3N}{2}}(2mE)^{\frac{3N}{2}-1}\frac{\frac{3N}{2}}{\Gamma(\frac{3N}{2}+1)} \tag{18}$$

(ν_E is an invariant positive measure, but it is not normalized.) The *microcanonical probability measure* is the normalized version of ν_E,

$$\mu_E(A) = \frac{\nu_E(A)}{\nu_E(\Gamma_E)} \tag{19}$$

which represents the "equal" probability distribution of all configurations $(\mathbf{q},\mathbf{p})$ of fixed energy E.

2 Derivation of the Maxwell Distribution

Now assuming the probability distribution μ_E on Γ_E, what is the resulting distribution of one particle?

We want to find the distribution function,

$$\mathbb{P}\left(\frac{1}{2m}p_1^2 \leq e\right) = \mu_E(A(e)) \tag{20}$$

where

$$A(e) = \{(\mathbf{q},\mathbf{p}) \in \Gamma_E | p_1^2 \leq 2me\} \tag{21}$$

Then the probability *density* function will be

$$f(e) = \frac{d}{de}\mathbb{P}\left(\frac{1}{2m}p_1^2 \leq e\right) = \frac{d}{de}\frac{d}{dE}\nu_E(A(e)(\Delta E)) \tag{22}$$

$$= \frac{L^3 4\pi 2me L^{3N-3}\nu_E^{3N-3}(\Gamma_{E-e}^{3N-3})}{\nu_E^{3N}(\Gamma_E^{3N})} \tag{23}$$

$$= 4\pi 2me\frac{(2m(E-e))^{\frac{3}{2}(N-1)-1}}{(2mE)^{\frac{3N}{2}-1}}\frac{\frac{3(N-1)}{2}}{\frac{3(N)}{2}}\frac{\Gamma(\frac{N}{2}+1)}{\Gamma(\frac{3(N-1)}{2}+1)} \tag{24}$$

$$\approx 4\pi 2meC\frac{1}{(2mE)^{3/2}}\left(1-\frac{e}{E})^{\frac{3N}{2}-\frac{5}{2}}\right)\left(\frac{3N}{2}\right)^{3/2} \tag{25}$$

where in the last (almost) equality, we have made use of Stirling's formula:

$$\Gamma(n) = \sqrt{2\pi} n^{n-1/2} \exp(-n) \left(1 - O(\frac{1}{n})\right) \tag{26}$$

We shall be interested in taking the *thermodynamic limit* in which we take

$$\begin{array}{ll} N \to \infty & \frac{N}{L^3} = \rho \\ L \to \infty & \frac{E}{L^3} = \eta \\ E \to \infty & \end{array} \tag{27}$$

Therefore, in the thermodynamic limit,

$$f(e) = C \frac{e}{\sqrt{2m}} \left(\frac{3\rho}{2\eta}\right)^{3/2} \exp(-\frac{3\rho}{2\eta} e) \tag{28}$$

Recalling now that $e = \frac{1}{2}mv^2$ where v is the speed of the first particle, we can change variables,

$$f(v) = Cv^2 \exp\left(-\beta \frac{1}{2} mv^2\right) \tag{29}$$

where $\beta = \frac{3\rho}{2\eta}$ is to be interpreted as inverse temperature.

3 Explanations and Issues

Why does the Maxwell distribution just derived so accurately describe the reality of molecules in a gas?

When we measure this distribution for a gas in equilibrium, we sample the system over a long time and by Poincare's recurrence theorem, explores the available portion of phase space over and over again such that the net effect is that the time averages become averages over phase space. To make this more precise, we will use Birkhoff's Ergodic theorem without proof.

Theorem 3.1. *Birkhoff*
For $\Gamma_0 \subset \Gamma$ *such that* $\mathrm{vol}(\Gamma_0) < \infty$ *and for all* $f \in L^1(\Gamma_0)$,

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f \circ \varphi_t(\boldsymbol{q}, \boldsymbol{p}) dt = \langle f \rangle_{(\boldsymbol{q},\boldsymbol{p})}$$

exists for almost all $(\boldsymbol{q}, \boldsymbol{p}) \in \Gamma_0$

Here $\langle f \rangle_{(\mathbf{q},\mathbf{p})}$ is the time average of f on a trajectory of dynamics with initial condition $(\mathbf{q}, \mathbf{p})$. The theorem says that this time average is defined almost everywhere on Γ_0. We can prove furthermore,

- $\langle f \rangle_{(\mathbf{q},\mathbf{p})}$ on a trajectory $(\mathbf{q}(t), \mathbf{p}(t))$ is actually independent of t

- If we assume further that Γ_0 is metrically indecomposable, then

$$\langle f \rangle_{(\mathbf{q},\mathbf{p})} = \frac{1}{\mathrm{vol}(\Gamma_0)} \int_{\Gamma_0} f(\mathbf{q},\mathbf{p}) d\mathbf{q} d\mathbf{p} \tag{30}$$

  independent of $(\mathbf{q},\mathbf{p}) \in \Gamma_0$ where the right hand side is the *phase space average* with respect to the invariant measure. Here, metric decomposability means that there is no non-trivial decomposition $\Gamma_0 = \Gamma_1 \cup \Gamma_2$ such that $\Gamma_1 \cap \Gamma_2 = \emptyset$ where Γ_1 and Γ_2 are of non-zero measure and are invariant under the flow.

So now one can argue that the microcanonical measure gives the correct result because phase space averages computed with it are equal to time averages. This makes sense except that the condition of metric indecomposability would have to be checked, which is usually impossibly difficult or else demonstrably false. In the case at hand, it is the latter, since Γ_E is actually metrically decomposable, since the energies of individual particles are conserved. Of course, in an experiment, we would have to pick many *different* particles to build a histogram of their energy distribution. So implicitly we are also averaging over a larger number of different particles. Whatever the precise explanation may be, the experimental fact is that the assumption of uniform distribution over a configuration of given energy gives the correct prediction under very general circumstances.

Statistical Mechanics, Math 266: Week 2 Notes

January 12 and 14, 2010

1 The Canonical Gibbs Measure

By analysis similar to the derivation of the Maxwell distribution, one can show for $\Lambda_0 \Subset \Lambda$, $N_0 \ll N$, $\frac{E}{\Lambda} = \eta$, $\frac{N}{\Lambda} = \rho$, and $N, E, \Lambda \to \infty$, the phase space distribution for Λ_0, and N_0 are obtained by conditioning $\mu_{E,\Lambda,N}$ on the N_0 particles found in the region Λ_0. In the thermodynamic limit,

$$\mu^{\text{Can}}_{\Lambda_0,N_0,\beta} \sim e^{-\beta H_{\Lambda_0,N_0}} \tag{1}$$

where $\beta = \frac{3\rho}{2\eta}$ as before. (See Minlos for details.) This is a special case of a canonical ensemble measure.

Note that the expected value of the energy of the subsystem in Λ_0 is then given by

$$\int H_{\Lambda_0,N_0}(\mathbf{q},\mathbf{p})\mu^{\text{Can}}_{\Lambda_0,N_0,\beta}(d\mathbf{q}d\mathbf{p}) = \frac{\int H_{\Lambda_0,N_0,\beta}(\mathbf{q},\mathbf{p})e^{-\beta H_{\Lambda_0,N_0,\beta}(\mathbf{qp})}d\mathbf{q}d\mathbf{p}}{Z^{\text{Can}}(\Lambda_0,N_0,\beta)} \tag{2}$$

Here, $Z^{\text{Can}}(\Lambda_0, N_0, \beta)$ is the partition function

$$Z^{\text{Can}}(\Lambda_0,N_0,\beta) = \int e^{-\beta H_{\Lambda_0,N_0}(\mathbf{q},\mathbf{p})}d\mathbf{q}d\mathbf{p} \tag{3}$$

Therefore

$$\langle H\rangle^{\text{Can}} = -\frac{d}{d\beta}\log Z^{\text{Can}}(\Lambda_0,N_0,\beta) \tag{4}$$

So we can compute Z^{Can} and derive the energy from this expression. For later use, we also define the free energy

$$F(\Lambda,N,\beta) = -\frac{1}{\beta}\log Z_{\Lambda,N,\beta} \tag{5}$$

To close out this section, we compute the partition function for a simple model. Let us suppose that we have N classical particles in a box $\Lambda \subset \mathbb{R}^3$ as before, with Hamiltonian given by $H(\mathbf{q},\mathbf{p}) = \sum_{i=1}^N \sum_{j\neq i} \frac{p_i^2}{2m} + U(|q_i - q_j|)$. If $U = 0$, then

$$Z_{\Lambda,N,\beta} = |\Lambda|^N \left(\int_{-\infty}^{\infty} e^{-\beta \frac{p^2}{2m}} dp \right)^{3N} \tag{6}$$

$$= \left(|\Lambda| \left(\frac{2\pi m}{\beta} \right)^{3/2} \right)^N \tag{7}$$

We can also calculate the expected energy

$$\langle H \rangle = -\frac{d}{d\beta} \log Z = \frac{3N}{2\beta} \tag{8}$$

So that we obtain equipartition of energy of $\frac{1}{2} k_B T$ per degree of freedom. In the case that $U \neq 0$, then $|\Lambda|^N$ is replaced by

$$\int_\Lambda dq_1 \ldots \int_\Lambda dq_N \prod_{1 \leq i < j \leq N} e^{-\beta U(|q_i - q_j|)}$$

2 The Ising Model

We shall be interested in the Ising Model because it is

- a simple model.
- historically significant in the development of Statistical Mechanics.
- a model that exhibits phenomena and allows us to ask questions that are central in Statistical Mechanics.
- a useful model for a number of physical phenomena, including some aspects of liquid-gas phase transitions, and adsorbtion of molecules on a surface such as in catalytic converter.

The point is that we will be able to understand the effects of interaction between particles in this model in considerable detail.

To define the model, we replace the continuous space $\mathbb{R}^d$ with the lattice $\mathbb{Z}^d$, and attach variables to each lattice site

$$\text{For all } x \in \mathbb{Z}^d \ , \ \sigma_x \in \{-1, +1\} \tag{9}$$

which can take only two values. One interpretation is that σ_x represents the rotation of a magnetic dipole attached to position x. Another useful interpretation is that of a "lattice gas" where we make the transformation

$$n_x = \frac{1}{2}(1 + \sigma_x) \tag{10}$$

$$n_x \in \{0, 1\} \tag{11}$$

so that each n_x is considered an occupation variable where a value of 0 denotes that the site x is empty, and a value of 1 denotes that site x is occupied by a gas molecule. One could think of a layer of gas molecules attaching to the surface of a catalytic material. In both cases, the model is only a caricature of the real stuff, but we will see that it is a useful caricature.

The phase space Γ is replaced by the configuration space

$$\begin{aligned} \Omega_\Lambda &= \{-1,+1\}^{|\Lambda|} && (12) \\ &= \{ \text{ all functions } \sigma : \Lambda \to \{-1,1\}\} && (13) \end{aligned}$$

where $\sigma = (\sigma_x)_{x\in\Lambda}$ is called a configuration. The minimal description of the physics that we need to do Statistical Mechanics is the energy, i.e., the Hamiltonian defined on Ω_Λ which assigns to each configuration its energy. (There are no Hamilton equations of motion here.)

$$H_\Lambda^{\text{Ising}} = -J \sum_{\substack{x,y\in\Lambda \\ |x-y|=1}} \sigma_x\sigma_y + h\sum_{x\in\Lambda}\sigma_x \tag{14}$$

One could generalize to $J(x,y)$ and $h(x)$ or even random variables J_{xy} but for now we restrict to constant J and h. J is a coupling constant and h gives the strength of an external magnetic field. We could again show from the microcanonical measure which assigns equal weight to all configurations of a given energy. With some effort we could then prove that restricted to the variables in a subvolume, and after taking $\Lambda \to \infty$, we get a canonical distribution. But we don't have to work with the subset of configuration of a given energy. In fact,

$$\mu_{\Lambda,\beta}(\sigma^\Lambda) = \frac{e^{-\beta H_\Lambda(\sigma^\Lambda)}}{Z(\Lambda,\beta)} \tag{15}$$

defines a probability measure in which lower energy configurations have higher probability. But counterbalancing this is the fact that there are many more configurations of energy $\sim \Lambda$ than there are of energy $= \min_\sigma H(\sigma)$. This will eventually lead us to introduce the concept of entropy and the competition between energy and entropy is at the core of the most interesting phenomena in Statistical Mechanics.

The Ising model is named after Ising, who studied it for his PhD thesis. It was his adviser, Lenz, who introduced it. It took a while to understand it, and Ising solved the one dimensional case in his dissertation. We will turn to that solution in a minute. It turned out to be less interesting than they had hoped, but Ising didn't realize that the dimension plays a crucial role.

What do we mean by "solution"? Here, we mean calculating Z, $\log Z$, or $\lim_{\Lambda\to\mathbb{Z}^d} \frac{1}{|\Lambda|}\log Z$. Let's first explain why this calculation indeed allows one to answer the most important questions. The free energy density in the thermodynamic limit is given by

$$f(\beta,h) = \lim_{\Lambda\nearrow\mathbb{Z}^d} -\frac{1}{\beta|\Lambda|}\log Z(\Lambda,\beta,h) \tag{16}$$

For a finite system $\Lambda \subset \mathbb{Z}^d$, the free energy is

$$F_\Lambda(\beta, h) = -\frac{1}{\beta} \log Z(\Lambda, \beta, h) \tag{17}$$

We already noted that

$$\langle H_\Lambda \rangle = \frac{\partial}{\partial \beta} \beta F_\Lambda(\beta, h) \tag{18}$$

Another important quantity is the magnetization (related to the particle number in the lattice gas interpretation)

$$\langle M_\Lambda \rangle = \frac{\sum_\sigma (\sum_{x \in \Lambda} \sigma_x) e^{-\beta H_\Lambda(\sigma)}}{Z_\Lambda(\beta, h)} \tag{19}$$

$$= -\frac{1}{\beta} \frac{\partial}{\partial h} \log Z_\Lambda(\beta, h) \tag{20}$$

$$= \frac{\partial}{\partial h} F_\Lambda(\beta, h) \tag{21}$$

For any "observable" A_Λ, one could add a term λA_Λ to H_Λ and define the corresponding $F_\Lambda(\beta, \lambda)$ so that

$$\langle A_\Lambda \rangle = \frac{\partial}{\partial \lambda} F_\Lambda(\beta, \lambda) \tag{22}$$

Moreover, it doesn't stop with the mean. We can also calculate variance and, in principle, any moment we like. For example,

$$V(A_\Lambda) = \langle (A_\Lambda - \langle A_\Lambda \rangle)^2 \rangle \tag{23}$$

$$= \langle A_\Lambda^2 \rangle - \langle A_\Lambda \rangle^2 \tag{24}$$

$$\text{and } \langle A_\Lambda^2 \rangle = \frac{\partial^2}{\partial \lambda^2} F_\Lambda(\beta, \lambda) \tag{25}$$

One can show that the limits

$$\lim_{\Lambda \nearrow \mathbb{Z}^d} \frac{1}{|\Lambda|} F_\Lambda(\beta, h) \tag{26}$$

$$\lim_{\Lambda \nearrow \mathbb{Z}^d} \frac{\langle H_\Lambda \rangle}{|\Lambda|} = e(\beta, h) = \frac{\partial}{\partial \beta} \beta f(\beta, h) \tag{27}$$

$$\lim_{\Lambda \nearrow \mathbb{Z}^d} \frac{\langle M_\Lambda \rangle}{|\Lambda|} = m(\beta, h) [= \frac{\partial}{\partial h} f(\beta, h)] \tag{28}$$

all exist, and where the energy density, e, was previously denoted η. Here interchanging the derivative and the thermodynamic limit is an interchange of limits and requires further justification. This interchange can be shown to be true under quite general assumptions, but this does not mean that all such limits are interchangeable. For example, $m(\beta, h)$ is, in general, *not* a continuous function of h and

$$\lim_{h \to 0^\pm} = \mp \mu \neq 0 \tag{29}$$

while $m(\beta, 0) = 0$. This is the signature of a phase transition. In fact, this is what Ising hoped to show. To this end, he calculated f and m for the one-dimensional Ising model.

3 More on the Ising Model

Recall that Hamiltonian of the the Ising model is

$$H_\Lambda = -J \sum_{\substack{x,y \in \Lambda \\ |x-y|=1}} \sigma_x \sigma_y \tag{30}$$

where we restrict to the situation where $J > 0$. This corresponds to the *ferromagnetic* Ising model. We have previously considered the thermodynamic limit as a limit in which a clear picture of the asymptotic behavior emerges. But there is more. In order to give a mathematically rigorous description of non-uniformity of the thermodynamic limit, the hallmark of phase transitions, we introduce the mathematical formulation of $*$-systems. We first focus on the Ising model and generalize later.

The set of configurations, Ω, is given by

$$\Omega = \prod_{x \in \mathbb{Z}} \{-1, +1\} \tag{31}$$

and we recall that Ω is compact with the Tychonoff topology. There is no difficulty with considering infinite configurations $\sigma \in \Omega$; in $d = 1$, σ is a bi-infinite sequence. In general, it just means that we are given a value $\sigma_x \in \{-1, +1\}$ for all $x \in \mathbb{Z}^d$.

It does not make sense in general to talk about probabilities of infinite configurations– not even ratios of probabilities, since typically "$\mu_\beta(\sigma) = 0$". The solutions is to limit the kind of events under consideration to the cylinder sets:

$$A_X(\sigma_X) = \{\eta \in \Omega | \eta_x = \sigma_x \text{ for all } x \in X\} \tag{32}$$

and $X \subset \mathbb{Z}^d$ is a *finite* set. Then we can make sense of $\mu(A_X)$ and the ratio

$$\frac{\mu(A_X(\sigma'_X))}{\mu(A_X(\sigma_X))} = e^{-\beta \Delta H_X(\sigma', \sigma)} \tag{33}$$

We can make this a bit more elegant by defining a fuction

$$f_A(\sigma) = \begin{cases} 1 & \text{if } \sigma \in A_X(\sigma_X) \\ 0 & \text{if } \sigma \notin A_X(\sigma_X \end{cases} \tag{34}$$

and considering linear combinations and uniform limits of such functions. This leads to the algebra of functions $C(\Omega)$, since such functions are continuous on Ω with the Tychonoff topology. Instead of defining a probability measure directly, we define the expectation of all such functions

$$\mu(f) = \sum_{\sigma \in \Omega} f(\sigma) \mu(\sigma) \tag{35}$$

By the Riesz Representation Theorem for measures, this is actually equivalent to defining a Borel probability measure on Ω.

References:

- Minlos
- Ruelle, Statistical Mechanics: Rigorous Results. London: Imperial College Press 1999
- H-O Georgii, Gibbs Measures and Phase Transitions. De Gruyter Studies in Mathematics Vol. 9. Berlin: de Gruyter 1988.

4 Explicit Calculation of the Thermodynamic Limit of the 1-Dimensional Ising Model

In the following section, we shall use the transfer matrix method which also appears in the analysis of transition matrices of Markov chains. For the Ising model in one dimension, we have

$$\Lambda = [-L, L] \subseteq \mathbb{Z} \tag{36}$$

$$H(\sigma) = -J \sum_{x=-L}^{L-1} \sigma_x \sigma_{x+1} \tag{37}$$

$$Z_\Lambda(\beta) = \sum_{\sigma_{-L},\ldots,\sigma_L=\pm1} e^{\beta J \sum_{x=-L}^{L-1} \sigma_x \sigma_{x+1}} \tag{38}$$

$$= \sum_{\sigma_{-L},\ldots,\sigma_L=\pm1} T_{\sigma_{-L}\sigma_{-L+1}} T_{\sigma_{-L+1}\sigma_{-L+2}} \cdots T_{\sigma_{L-1}\sigma_L} \tag{39}$$

where we have

$$T = \begin{pmatrix} T_{++} & T_{+-} \\ T_{-+} & T_{--} \end{pmatrix} = \begin{pmatrix} e^{\beta J} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta J} \end{pmatrix} \tag{40}$$

It follows then that in terms of the transfer matrix, T,

$$\left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, T^{2L-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle \tag{41}$$

which can be simplified further, since T is clearly a diagonalizable matrix. The eigenvalues can clearly be seen to be:

$$\lambda_+ = 2\cosh(\beta J) \tag{42}$$

$$\lambda_- = 2\sinh(\beta J) \tag{43}$$

Thus,

$$T = S \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix} S^{-1} \tag{44}$$

$$Z_\Lambda(\beta) = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, S \begin{pmatrix} \lambda_+^{2L-1} & 0 \\ 0 & \lambda_-^{2L-1} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle = 2(2\cosh(\beta J))^{2L-1} \tag{45}$$

Then by taking the thermodynamic limit $\Lambda \to \mathbb{Z}$, we obtain the free energy density,

$$f(\beta) = -\frac{1}{\beta}\log(2\cosh(\beta J)) \tag{46}$$

in a similar way, we can compute the expectation

$$\mu_\Lambda(A_{[1,n]})(\sigma_1, \ldots, \sigma_n) = \frac{\sum_{\substack{\sigma_{-L},\ldots,\sigma_0=\pm 1 \\ \sigma_{n+1},\ldots,\sigma_L=\pm 1}} e^{-\beta H(\sigma)}}{Z_\Lambda(\beta)} \tag{47}$$

where we hold $\sigma_1, \ldots, \sigma_n$ fixed. Expanding this now using the transfer matrix method, we obtain for the numerator

$$\begin{aligned} &\sum_{\sigma_{-L}\sigma_L=\pm 1} (T^L)_{\sigma_{-L}\sigma_1} T_{\sigma_1\sigma_2} T_{\sigma_2\sigma_3} \ldots T_{\sigma_{n-1}\sigma_n} (T^{L-n-1})_{\sigma_n\sigma_L} \\ &= e^{-\beta H_{[1,n]}(\sigma_1,\ldots,\sigma_n)} \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, T^L e_{\sigma_1} \right\rangle \left\langle e_{\sigma_n}, T^{L-n-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle \end{aligned}$$

where we use the notation

$$e_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$
$$e_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

And so the final result is

$$\mu_{[-L,L]}(A_{[1,n]}(\sigma_1, \ldots, \sigma_n)) = \frac{e^{-\beta H(\sigma_1,\ldots,\sigma_n)} \lambda_+^L \sigma_+^{L-n-1}}{2\lambda_+^{2L-1}} \tag{48}$$

$$= \frac{e^{-\beta H(\sigma_1,\ldots,\sigma_n)}}{2\lambda_+^{n-1}} \tag{49}$$

Statistical Mechanics, Math 266: Week 3 Notes

January 19 and 21, 2010

1 Phase Transitions and Spontaneous Symmetry Breaking

Consider the d-dimensional Ising model:

$$\Lambda \subseteq \mathbb{Z}^d \text{ e.g., } \Lambda = [1, L]^d \tag{1}$$

$$H_\Lambda = -J \sum_{\substack{|x-y|=1 \\ x,y\in\Lambda}} \sigma_x \sigma_y \tag{2}$$

As before, we will assume that the model is ferromagnetic, so $J > 0$. The Hamiltonian H_Λ exhibits a spin flip symmetry which takes $\sigma_x \to -\sigma_x$ Many fundamental models have symmetries and many interesting phase transitions are accompanied by symmetry breaking. More precisely, let $F : \Omega \to \Omega$ be defined by

$$F(\eta) = -\eta \tag{3}$$

Clearly $H_\Lambda \circ F = H_\Lambda$. It follows that the equilibrium state

$$\omega_\Lambda(f) = \frac{\sum_\eta f(\eta) e^{-\beta H_\Lambda(\eta)}}{\sum_\eta e^{-\beta H_\Lambda(\eta)}} \tag{4}$$

is also F-symmetric, meaning

$$\omega_\Lambda(f \circ F) = \omega_\Lambda(f) \text{ for all } f \in C(\Omega_\Lambda) \tag{5}$$

In particular, we have for all $x \in \Lambda$,

$$\omega_\Lambda(\sigma_x) = \omega_\Lambda(\sigma_x \circ F) = -\omega_\Lambda(\sigma_x) \tag{6}$$

$$\text{and hence } \omega_\Lambda(\sigma_x) = 0 \text{ for all } x \in \Lambda \tag{7}$$

Taking the thermodynamic limit does not change this,

$$\lim_{\Lambda \nearrow \mathbb{Z}^d} \omega_\Lambda(\sigma_x) = 0 \tag{8}$$

All of the thermodynamics is contained in the function $f(\beta)$, the free energy density

$$-\beta f(\beta) = \lim_{\Lambda \nearrow \mathbb{Z}^d} \frac{1}{|\Lambda|} \log Z_\Lambda(\beta) \tag{9}$$

$$Z_\Lambda(\beta) = \sum_{\eta \in \Omega_\Lambda} e^{-\beta H_\Lambda(\eta)} \tag{10}$$

It is easy to see that boundary conditions do not affect f, so in the thermodynamic limit, we obtain the same thermodynamics.

Now define the boundary of Λ,

$$\partial\Lambda = \{x \in \Lambda | \text{ there exists } y \in \mathbb{Z}^d \cup \Lambda^c \text{ and } |x-y| = 1\} \tag{11}$$

Then consider $b_\Lambda \in C(\Omega_{\partial\Lambda})$ and suppose that $\|b_\Lambda\|_{\sup} \leq B|\partial\Lambda|$, so that b_Λ is uniformly bounded with B some fixed constant. For some sequence of boundary terms, we may find that

$$\omega^b = \lim_\Lambda \omega^b_\Lambda \text{ exists} \tag{12}$$

$$\omega^{b_\Lambda}_\Lambda(f) = \frac{1}{Z^{b_\Lambda}_\Lambda(\beta)} \sum_{\eta \in \Omega_\Lambda} f(\eta) e^{-\beta H^{b_\Lambda}_\Lambda(\eta)} \tag{13}$$

$$\text{where } H^{b_\Lambda}_\Lambda = H_\Lambda + b_\Lambda \tag{14}$$

Then,

$$e^{-B\beta|\partial\Lambda|} Z_\Lambda \leq Z^{b_\Lambda}_\Lambda \leq Z_\Lambda e^{B\beta|\partial\Lambda|} \tag{15}$$

So that as long as we have $\frac{|\partial\Lambda|}{|\Lambda|} \to 0$, we will have that $f^b(\beta) = f(\beta)$. It is therefore reasonable to assume that $\omega^b = \lim_\Lambda \omega^{b_\Lambda}_\Lambda$. If this limit exists, it will also describe the equilibrium of the thermodynamic system. (We will make this precise and rigorous when we study characteristics of equilibrium later.) Under quite general conditions, one can show that for some $\beta_c > 0$, ω^b is independent of b for all $0 \leq \beta \leq \beta_c$. But it often happens that there is some dependence on the boundary condition, b, if β is large enough. Before doing anything more general, we will show that this happens for the d-dimensional Ising model.

2 The Peierls Argument

We will consider the particular boundary term leading to what is called $+$ boundary conditions,

$$b_\Lambda = -J \sum_{\substack{x \in \partial\Lambda \\ y \in \Lambda^c, |x-y|=1}} \sigma_x \cdot 1 \tag{16}$$

as if the spin at $y \in \Lambda^c$ are all fixed to be $+1$. $-$ boundary conditions are completely analogous. Let's assume that the following limit exists:

$$\omega^+ = \lim_\Lambda \omega^+_\Lambda \tag{17}$$

$$= \lim_\Lambda \omega^{b_\Lambda}_\Lambda \tag{18}$$

Note that if we were free to interchange limits, it would be rather trivial to show that $\lim_{\beta\to\infty} m(\beta) = 1$ since $\lim_{\beta\to\infty} \omega_\Lambda^+(\sigma_x) = 1$ for all finite $\Lambda \subseteq \mathbb{Z}^d$, for all $d \geq 1$.

Theorem 2.1. *Let $J > 0$ and $d = 2$. Then,*

1. *There exists $\beta_1 > 0$ such that for all $\beta > \beta_1$,*

$$\frac{1}{|\Lambda|}\sum_{x\in\Lambda} \omega_\Lambda^+(\sigma_x) = m_\Lambda(\beta) > 0 \tag{19}$$

and $\lim_{\beta\to\infty} m_\Lambda(\beta) = 1$ uniformly in Λ.

2. *There exists $\beta_2 > 0$ such that for all $\beta > \beta_2$ and for all $x \in \Lambda$,*

$$\omega_\Lambda^+(\sigma_x) = m_x(\beta) > 0 \tag{20}$$

and $\lim_{\beta\to\infty} m_x(\beta) = 1$ uniformly in x.

In fact, we obtain bounds of the form

$$0 \leq 1 - m_x(\beta) \leq 216e^{-8J\beta} \tag{21}$$

$$0 \leq 1 - m_\Lambda(\beta) \leq 216e^{-8J\beta} \tag{22}$$

For sufficiently large β.

Remark 2.1. *Onsager obtained an exact solution of the free energy density if the 2-dimensional Ising model frem which it follows that $\beta_c = \frac{\log(1+\sqrt{2})}{2J}$.*

To prove the theorem we will use the contour description of

$$\Omega_\Lambda^+ = \{\eta \in \Omega_{\Lambda\cup\partial(\Lambda^c)} | \eta \restriction_{\Lambda^c} = +1\} \tag{23}$$

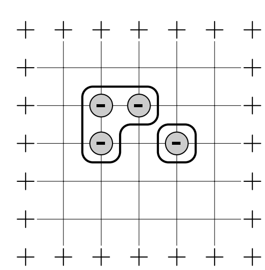

Figure 1: Configuration space as described by contours.

There is a one-to-one correspondence between configurations and configurations of contours.

$$\sigma \in \Omega_\Lambda^+ \longleftrightarrow \{\gamma_1, \ldots, \gamma_n\} = \Gamma(\sigma) \tag{24}$$

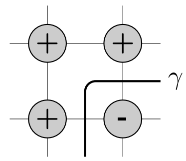

Figure 2: A dual edge is in γ if the corresponding edge has a $+-$ pair.

Here, $\Gamma(\sigma)$ is a configuration of closed, non-intersecting paths (where we ignore corner intersection) in the dual lattice. We define the support of a configuration,

$$\operatorname{supp}\Gamma = \bigcup_{i=1}^{n} \operatorname{supp}\gamma_i = \{(x,y) \in \mathbb{Z}^2 | \sigma_x\sigma_y = -1\} \tag{25}$$

and this allows us to define the length of a configuration of contours as

$$\ell(\Gamma) = |\operatorname{supp}\Gamma| = \sum_{i=1}^{n} \ell(\gamma_i) \tag{26}$$

Now, we wish to rewrite the Hamiltonian of the Ising model in terms of contours.

$$H_\Lambda^+(\sigma) = -J\{\#\{(x,y) \in \mathbb{Z}^2 | \sigma_x\sigma_y = +1\}\} + J\{\#\{(x,y) \in \mathbb{Z}^2 | \sigma_x\sigma_y = -1\}\} \tag{27}$$

$$= -J|B(\Lambda)| + 2J\ell(\Gamma) \tag{28}$$

where $|B(\Lambda)|$ is equal to the number of edges in Λ. Here, we can think of energy as being proportional to the length of the contours. Since the configurations $\Gamma(\sigma)$ that we consider consist of compatible, closed, non-intersecting paths, we define $V(\gamma)$ to be the vertices enclosed by a given contour γ.

Lemma 2.1. *For all γ such that $\ell(\gamma) < \infty$,*

$$|V(\gamma)| \leq \frac{1}{16}\ell(\gamma)^2 \tag{29}$$

Proof. Left to reader. See homework. □

Lemma 2.2. *Let $\Lambda \subseteq \mathbb{Z}^2$ and $\ell = 4, 6, 8, \ldots$ Define $M_\Lambda(\ell) = \#$ of distinct simple contours of length ℓ within Λ (with '+' boundary conditions). Then*

$$M_\Lambda(\ell) \leq 3^{\ell-1}|\Lambda| \tag{30}$$

Proof. Left to reader. See homework. □

We know what $\mathbb{P}_\Lambda^+(\sigma)$ is, and this allows us to make sense of $\mathbb{P}(\Gamma)$. Now, for a given contour γ, we define $\mathbb{P}_\Lambda^+(\gamma)$ to be the probability that γ occurs. Specifically, it is the event that contains all configurations Γ that have γ in it. Explicitly,

$$\mathbb{P}_\Lambda^+(\gamma) = \frac{\sum_{\substack{\sigma\in\Omega_\Lambda^+ \\ \gamma\in\Gamma(\sigma)}} e^{-\beta H_\Lambda^+(\sigma)}}{\sum_{\sigma\in\Omega_\Lambda^+} e^{-\beta H_\Lambda^+(\sigma)}} \tag{31}$$

Lemma 2.3 (Peierls Estimate)**.**

$$\mathbb{P}_\Lambda^+(\gamma) \le e^{-2J\beta\ell(\gamma)} \tag{32}$$

Proof. For all σ such that $\gamma \in \Gamma(\sigma)$, for some fixed γ define σ^* as the unique configuration such that $\Gamma(\sigma^*) = \Gamma(\sigma)\backslash\{\gamma\}$. Explicitly, σ^* is obtained by flipping all spins located at $x \in V(\gamma)$. Recall that $H_\Lambda^+(\sigma) = -JB(\Lambda) + 2J\ell(\gamma)$, and therefore

$$H_\Lambda^+(\sigma) - H_\Lambda^+(\sigma^*) = 2J\ell(\Gamma) \tag{33}$$

and

$$\begin{aligned}
\mathbb{P}_\Lambda^+(\gamma) &= \frac{\sum_{\substack{\sigma\in\Omega_\Lambda^+ \\ \gamma\in\Gamma(\sigma)}} e^{-\beta H_\Lambda^+(\sigma)}}{\sum_{\sigma\in\Omega_\Lambda^+} e^{-\beta H_\Lambda^+(\sigma)}} && (34)\\
&\le \frac{\sum_{\substack{\sigma\in\Omega_\Lambda^+ \\ \gamma\in\Gamma(\sigma)}} e^{-\beta H_\Lambda^+(\sigma)}}{\sum_{\substack{\sigma\in\Omega_\Lambda^+ \\ \gamma\in\Gamma(\sigma)}} e^{-\beta H_\Lambda^+(\sigma^*)}} && (35)\\
&\le e^{-2\beta J\ell(\gamma)} && (36)
\end{aligned}$$

□

Proof of Theorem 2.1. We will estimate

$$0 \le 1 - \omega_\Lambda^+(\sigma_x) = \omega_\Lambda^+(1 - \sigma_x) \tag{37}$$

Observe that $1 - \sigma_x$ takes the values 0 and 2. If $1 - \sigma_x = 2$, then there exists a $\gamma \in \Gamma(\eta)$ such that $x \in V(\gamma)$. Denote by $\gamma^*(\sigma)$ the *first* contour you meet starting at x. Then

$$1-\omega_\Lambda^+(\sigma_x) \leq \frac{2\sum_{\gamma,x\in V(\Gamma)}\sum_{\sigma,\gamma^*(\sigma)=\gamma} e^{-\beta H_\Lambda^+(\sigma)}}{\sum_\sigma e^{-\beta H_\Lambda^+(\sigma)}} \tag{38}$$

$$\leq 2\sum_{\gamma,x\in V(\gamma)} \frac{\sum_{\sigma,\gamma\in\Gamma(\sigma)} e^{-\beta H_\Lambda^+(\sigma)}}{\sum_\sigma e^{-\beta H_\Lambda^+(\sigma)}} \tag{39}$$

$$= 2\sum_{\gamma,x\in V(\gamma)} \mathbb{P}_\Lambda^+(\gamma) \tag{40}$$

$$\leq 2\sum_{\gamma,x\in V(\gamma)} e^{-2\beta J\ell(\gamma)} \tag{41}$$

where (39) comes from the fact that $1-\omega_\Lambda^+$ vanishes if it is enclosed by an even number of contours, and (41) follows from Lemma 2.3. From here, Parts 1 and 2 of Theorem 2.1 proceed only slightly differently, and so we only present the proof of Part 2. We rewrite the inequality (41) in a more suggestive way.

$$1-\omega_\Lambda^+(\sigma_x) \leq 2\sum_{\ell=4,6,8,\ldots}\sum_{\substack{\gamma,x\in V(\gamma)\\ \ell(\gamma)=\ell}} e^{-2\beta J\ell} \tag{42}$$

$$\leq 2\sum_{\ell=4,6,8,\ldots} \ell^2 3^\ell e^{-2\beta J\ell} \tag{43}$$

$$\leq 16(3e^{-2\beta J})^2 \text{ where } 3e^{-2\beta J} \leq \frac{1}{2} \tag{44}$$

Here, we have used the observation that a contour of length ℓ must be contained within a square box of size ℓ centered at x, along with lemma 2.2 and the fact that

$$\sum_{k=2}^{\infty} k^2 r^k = \frac{2r^2(2-\frac{3}{2}r+\frac{1}{2}r^2)}{(1-r)^3} \tag{45}$$

The lower bound on β now follows. $\square$

Remark 2.2. *Clearly $\omega_\Lambda^+(\sigma_x) \to 1$ as $\beta\to\infty$, and with the same estimates for the $-$ boundary condition, we have that $\omega_\Lambda^-(\sigma_x) \to -1$ as $\beta\to\infty$. Hence, if*

$$\omega_\Lambda^\pm \to \omega^\pm \text{ as } \Lambda \nearrow \mathbb{Z}^2 \tag{46}$$

clearly, $\omega^+ \neq \omega^-$. A similar argument works for $d \geq 2$, and these same arguments (using Peierls Estimate) can be generalized to other models with somewhat similar structure.

3 The Griffiths Inequalities

The goal for us is to show the existence of the limiting Gibbs states, but the Griffiths inequalities have many other applications. Again, we are considering

Ising systems on $\mathbb{Z}^d$. The algebra of observables for a finite volume $\Lambda \subset \mathbb{Z}^d$ is $C(\Omega_\Lambda)$. Consider the special observables

$$\sigma_A = \prod_{x \in A} \sigma_x \text{ for all } A \subseteq \Lambda \tag{47}$$

$$\sigma_\emptyset = 1 \tag{48}$$

We make the observation that the set $\{\sigma_A | A \subseteq \Lambda\}$ is a basis for $C(\Omega_\Lambda)$ because $\delta_{\eta_x = \varepsilon} = \frac{1}{2}(1 + \varepsilon \sigma_x)$ forms a basis upon taking products. The ferromagnetic Ising model can be generalized to a general class of ferromagnetic models with local Hamiltonians of the form

$$H_\Lambda = -\sum_A J_A \sigma_A \ , \ J_A \geq 0 \tag{49}$$

Note that

$$\frac{\partial}{\partial J_B} \omega_\Lambda(\sigma_A) = \omega_\Lambda(\sigma_A \sigma_B) - \omega_\Lambda(\sigma_A)\omega_\Lambda(\sigma_B) \tag{50}$$

where we have set $\beta = 1$ in this equation.

Theorem 3.1 (Griffiths Inequalities)**.** *Let ω_Λ be the Gibbs state at β with ferromagnetic Hamiltonian H_Λ. Then*

1. $\omega_\Lambda(\sigma_A) \geq 0$ *for all* $A \subset \Lambda$.
2. $\omega_\Lambda(\sigma_A \sigma_B) - \omega_\Lambda(\sigma_A)\omega_\Lambda(\sigma_b) \geq 0$ *for all* $A, B \subset \Lambda$.

Proof. Proof of 1.

$$\omega_\Lambda(\sigma_A) = \frac{1}{Z_\Lambda} \sum_{\eta \in \Omega_\Lambda} \sigma_A(\eta) e^{-\beta H_\Lambda(\eta)} \tag{51}$$

$$= \frac{1}{Z_\Lambda} \sum_{n=0}^{\infty} \frac{\beta^n}{n!} \sum_{\eta \in \Omega_\Lambda} \sigma_A(\eta) (\sum_{B \subseteq \Lambda} J_B \sigma_B)^n \tag{52}$$

Clearly $\sigma_A \sigma_B = \sigma_C$ with $C = A \triangle B$, and where $A \triangle B = (A \cup B) \setminus (A \cap B)$ is the symmetric difference of A and B. Now, we group all terms with the same C.

$$\omega_\Lambda(\sigma_A) = \sum_C a(C) \sum_{\eta \in \Omega_\Lambda} \sigma_C(\eta) \tag{53}$$

If $C \neq \emptyset$, $\sum_{\eta \in \Omega_\Lambda} \sigma_C(\eta) = 0$ since if $x \in C$ then the sum over η with $\eta_x = \pm 1$ cancel each other out. In the case that $C = \emptyset$, $\sum_{\eta \in \Omega_\Lambda} \sigma_C = 2^{|\Lambda|}$. So $\omega_\Lambda(\sigma_A) = a(\emptyset) 2^{|\Lambda|} \geq 0$.

Proof of 2.

Note that for 1, we did not really use the structure of $\mathbb{Z}^d$, and the argument works on any finite set Λ. Now we will consider $\widetilde{\Lambda} = \Lambda \sqcup \Lambda$, two disjoint copies of Λ. Equivalently, we can consider a system with two copies of the algebra

$$\widetilde{\mathcal{A}_\Lambda} = C(\Omega_\Lambda) \otimes C(\Omega_\Lambda) \tag{54}$$

where each copy of $C(\Omega_\Lambda)$ is generated by the functions σ_x and τ_x respectively. The configuration space is $\widetilde{\Omega_\Lambda} = \Omega_\Lambda \times \Omega_\Lambda = \{(\eta, \xi) | \eta_x, \xi_x \in \{-1, +1\}\}$. Similarly, we have σ_A and τ_A. Define

$$\widetilde{H_\Lambda}(\eta, \xi) = H_\Lambda(\eta) + H_\Lambda(\xi) \tag{55}$$

$$\widetilde{Z_\Lambda} = \sum_{\eta \in \Omega_\Lambda} \sum_{\xi \in \Omega_\Lambda} e^{-\beta H_\Lambda(\eta)} e^{-\beta H_\Lambda(\xi)} \tag{56}$$

$$= (Z_\Lambda)^2 \tag{57}$$

If $f(\eta, \xi) = f_1(\eta) f_2(\xi)$, we have

$$\widetilde{\omega_\Lambda}(f) = \omega_\Lambda(f_1)\omega_\Lambda(f_2) \tag{58}$$

Now consider "rotated" variables

$$s_x = \frac{1}{\sqrt{2}}(\sigma_x + \tau_x) \tag{59}$$

$$t_x = \frac{1}{\sqrt{2}}(\sigma_x - \tau_x) \tag{60}$$

which take values $-\sqrt{2}, 0, \sqrt{2}$ on double configurations. Note that

$$\sigma_x = \frac{1}{\sqrt{2}}(s_x + t_x) \tag{61}$$

$$\tau_x = \frac{1}{\sqrt{2}}(s_x - t_x) \tag{62}$$

and for $A \subseteq \Lambda$,

$$\Delta_A^{\pm} = \sigma_A \pm \tau_A = \left(\frac{1}{\sqrt{2}}\right)^{|A|} \{(s+t)_A \pm (s-t)_A\} \tag{63}$$

where $\sigma_A = \prod_{x \in A} \sigma_x$, and τ_A is completely analogous.

Lemma 3.1.

$$\Delta_A^{\pm} = \sum_{B \subseteq A} K_B s_{A \setminus B} t_B \tag{64}$$

with some $K_B \geq 0$

Proof. Just calculate

$$(s+t)_A = \prod_{x \in A} (s_x + t_x) \tag{65}$$

$$= \sum_{B \subseteq A} s_{A \setminus B} t_B \tag{66}$$

and

$$(s-t)_A = \sum_{B \subseteq A} (-1)^{|B|} s_{A \setminus B} t_B \tag{67}$$

The lemma follows. □

Now, we can proceed with the proof of 2.

$$\begin{aligned}
\omega_\Lambda(\sigma_A\sigma_B) - &\omega_\Lambda(\sigma_A)\omega_\Lambda(\sigma_B) && (68)\\
&= \widetilde{\omega_\Lambda}(\sigma_A\sigma_B) - \widetilde{\omega_\Lambda}(\sigma_A\tau_B) && (69)\\
&= \widetilde{\omega_\Lambda}(\sigma_A(\sigma_B - \tau_B)) && (70)\\
&= \left(\frac{1}{\sqrt{2}}\right)^{|A|+|B|} \widetilde{\omega_\Lambda}(s+t)_A\{(s+t)_B + (s-t)_B\} && (71)\\
&= \widetilde{\omega_\Lambda}(\sum_{C\subseteq B} K_C (s+t)_A s_{B\setminus C} t_C) && (72)\\
&= \widetilde{\omega_\Lambda}(\sum_{D} K_D s_{A\cup B\setminus D} t_D) && (73)
\end{aligned}$$

The variables s_x and t_x have the following properties:

$$s_x t_x = 0 \text{ and either } s_x \text{ or } t_x = 0 \text{ for a given configuration } \sigma_x, \tau_x \qquad (74)$$

t_x and s_x are odd functions of σ_x and τ_x, so all their odd powers are odd and all their even porwers are of course greater than or equal to 0. So

$$\sum_{\sigma_x,\tau_x} s_x^n = \begin{cases} = 0 & \text{if } n \text{ odd.} \\ > 0 & \text{if } n \text{ even.} \end{cases} \qquad (75)$$

and the same result holds for $\sum t_x^n$. The Hamiltonian can be rewritten,

$$\begin{aligned}
\widetilde{H_\Lambda} &= \sum_{A\subset\Lambda} K_A\sigma_A + K_A\tau_A && (76)\\
&= \sum_{A\subset\Lambda} \widetilde{K_A} \sum_{C\subset A} (1+(-1)^{|C|}) s_{A\setminus C} t_C && (77)
\end{aligned}$$

and is again a Hamiltonian with coefficients greater than or equal to 0 in the monomial basis. Although the polynomials in s and t are not independent variables, the same argument as in 1 applies.

$$\begin{aligned}
\widetilde{H_\Lambda} &= \sum_{C,D\subset\Lambda} \widetilde{K_{C,D}} s_C t_D && (78)\\
\widetilde{\omega_\Lambda} &= \frac{1}{\widetilde{Z_\Lambda}} \sum_{n=0}^{\infty} \frac{\beta^n}{n!} \sum_{\{\sigma_x,\tau_x\}} s_A t_B (\sum_{C,D\subset A} \widetilde{K_{C,D}} s_C t_D)^n && (79)\\
&= \frac{1}{\widetilde{Z_\Lambda}} \sum_{n=0}^{\infty} \frac{\beta^n}{n!} \sum_{\{\sigma_x,\tau_x\}} \sum_{C,D} \widetilde{\widetilde{K_{ABCD}}} s_A s_C t_B t_D && (80)
\end{aligned}$$

Indeed,

$$\begin{aligned}
\widetilde{H_\Lambda} &= \sum_{A\subset\Lambda} J_A\sigma_A + J_A\tau_A && (81)\\
&= \frac{1}{\sqrt{2}} \sum_{A\subset\Lambda} \sum_{C\subset A} J_A (1+(-1)^{|C|}) s_{A\setminus C} t_C && (82)
\end{aligned}$$

Therefore we can apply 1 to $\widetilde{H_\Lambda}$ and finish the proof of 2. $\square$

Recall the observation

$$\frac{\partial}{\partial J_B}\omega_\Lambda(\sigma_A) = \omega_\Lambda(\sigma_A\sigma_B) - \omega_\Lambda(\sigma_A)\omega_\Lambda(\sigma_B) \tag{83}$$

therefore the second Griffiths inequality implies that $\frac{\partial}{\partial J_B}\omega_\Lambda(\sigma_A) \geq 0$ for ferromagnetic Ising models.

4 The Thermodynamic Limit of Ising Equilibrium States

The set of states on $C(\Omega_{\mathbb{Z}^d})$ is weak-* compact. From this we deduce that at each fixed βJ, the set of finite volume states $\{\omega_\Lambda\}_{\Lambda\subseteq\mathbb{Z}^d}$ has at least one limit point (extend them to states on all of $\mathbb{Z}^d$ in more or less any way you like). And more generally the same is true for sequences with other boundary conditions b_Λ. The Peierls argument shows that if ω^+ and ω^- are such limit points of $\{\omega_\Lambda^+\}$ and $\{\omega_\Lambda^-\}$ respectively, then for all β large enough, they will be distinct, since

$$\omega^+(\sigma_0) = -\omega^-(\sigma_0) \neq 0 \tag{84}$$

Later, in a more general context, we will show that for small β the limit points are unique independent of b_Λ. It is nevertheless still an interesting question whether the sequence ω_Λ^+ itself converges.

Theorem 4.1. *1. Let $\{\omega_\Lambda^0\}$ be the sequence of β Gibbs states in finite volume $\Lambda \subseteq \mathbb{Z}^d$ of the Ising model with free boundary conditions. Then*

$$\omega_\Lambda^0(\sigma_A) \nearrow \omega^0(\sigma_A) \text{ for all finite subsets } A \subset \mathbb{Z}^d \tag{85}$$

2. If $\{\omega_\Lambda^+\}$ is a sequence corresponding to + boundary conditions, then

$$\omega_\Lambda^+(\sigma_A) \searrow \omega^+(\sigma_A) \tag{86}$$

i.e., we have weak convergence in both cases and they are monotone increasing and decreasing, respectively, on the basis functions σ_A.*

Proof. Proof of 1.

ω_Λ^0 can be regarded as the Gibbs state for the ferromagnetic Ising model

$$H_\Lambda = -\sum_X J_X^\Lambda \sigma_X \tag{87}$$

with

$$J_X^\Lambda = \begin{cases} J & \text{if } X = (x,y), x,y \in \Lambda, |x-y| = 1 \\ 0 & \text{otherwise} \end{cases} \tag{88}$$

with $J > 0$. Hence, J_X^Λ is monotonic increasing by the second Griffiths inequality.

Proof of 2.

Define

$$J_X^\Lambda = \begin{cases} J_x & \text{if } X = x, y, x, y \in \Lambda, |x-y| = 1 \\ +\infty & \text{if } X = \{x\}, x \in \Lambda^c \\ 0 & \text{otherwise} \end{cases} \tag{89}$$

It is not hard to see that the infinite coupling constant does not pose a problem. $\square$

Note that the Griffiths inequalities also show that a variety of other Ising models with higher dimensionality and anisotropies also have a non-vanishing magnetism at sufficiently low temperature by comparing with the two-dimensional translation invariant model.

Statistical Mechanics, Math 266: Week 4 Notes

January 26 and 28, 2010

1 Quantum Statistical Mechanics

The canonical formalism of Quantum Statistical Mechanics has the same structure as in the classical case, but the mathematical objects playing the role of the probability measure and the Hamiltonian are of a different kind. We will first set up the mathematical structure, then make it more explicit in the context of Quantum Spin Systems. Later, we will also consider systems of quantum particles in the continuum. A *finite* quantum system, analogous to a finite number of classical particles in a finite volume $\Lambda \subseteq \mathbb{R}^3$, or the Ising model with a finite number of spins, is described by

- a separable Hilbert space of states $\mathcal{H}$
- a densely defined self-adjoint operator, H on $\mathcal{H}$, which plays the role of the Hamiltonian
- observables described by bounded operators on $\mathcal{H}$, usually a norm-closed *-subalgebra of $\mathcal{B}(\mathcal{H})$

Example 1.1. N non-interacting, spinless, structureless, distinguishable point particles in a box $\Lambda \subseteq \mathbb{R}^d$:

$$\mathcal{H} = L^2(\Lambda^N, dx) \tag{1}$$

$$H_0 = -\frac{\hbar}{2m} \sum_{i=1}^{N} \Delta_i \tag{2}$$

$$D(H) = H^2 \tag{3}$$

where Δ_i is the d-dimensional Laplacian with suitable boundary conditions, and H^2 is the Sobolev Space of twice weakly differentiable functions in L^2.

Example 1.2. N spinless, structureless point particles interacting via a pair potential V in $\Lambda \subseteq \mathbb{R}^d$.

$$H = H_0 + \sum_{1 \le i < j \le N} V(x_i - x_j) \tag{4}$$

Example 1.3. N non-interacting spin $\frac{1}{2}$ Fermions.

$$\mathcal{H} = [L^2(\Lambda) \otimes \mathbb{C}^2]^{\wedge N} \tag{5}$$

$$H = H_0 \otimes \mathbb{1} \tag{6}$$

where H_0 is the Hamiltonian defined in Example 1.1.

Example 1.4. Two-state quantum spins or qubits on labelled sites $x \in \Lambda \subseteq \mathbb{Z}^d$

$$\mathcal{H} = (\mathbb{C}^2)^{\otimes |\Lambda|} \tag{7}$$

One typical model is the Heisenberg model with the Hamiltonian

$$H = -J \sum_{\substack{x,y \in \Lambda \\ |x-y|=1}} \vec{S}_x \cdot \vec{S}_y \tag{8}$$

2 The Canonical Formalism for Quantum Systems

We will assume (and not encounter exceptions to this assumption) that the finite system Hamiltonian H_Λ is such that $e^{-\beta H_\Lambda}$ is trace class for all $\beta > 0$. The index Λ indicates that typically we will consider again sequences of systems indexed by finite volume Λ which will be taken to $\mathbb{R}^d$, $\mathbb{Z}^d$, or whatever. For self-adjoint H, the condition that $e^{-\beta H}$ is equivalent to the statement that its spectrum consists entirely of eigenvalues of finite multiplicity, which we will often enumerate as follows:

$$\lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \dots \tag{9}$$

repeated according to their multiplicity, and the assumption that

$$\sum_{n \geq 0} e^{-\beta \lambda_n} < +\infty \tag{10}$$

If $\dim \mathcal{H}_\Lambda < +\infty$, this is of course automatically satisfied. For systems of particles in finite volume, there are general theorems about operators of the form

$$H = \frac{1}{2m} \sum_{i=1}^{N} \Delta_i + \sum_{1 \leq i < j \leq N} V(x_i - x_j) \tag{11}$$

that guarantee this property for most systems of interest. Since $e^{-\beta H}$ is assumed to be trace class, we can define

$$\rho_\beta = \frac{1}{Z} e^{-\beta H} \tag{12}$$

where $Z = \operatorname{Tr} e^{-\beta H} = \sum_{n \geq 0} e^{-\beta \lambda_n}$. ρ_β is then a positive definite operator of trace class and with trace 1:

$$\operatorname{Tr} \rho_\beta = 1 \tag{13}$$

This is what is called a density matrix.

Density matrices in quantum mechanics play the role of probability densities for classical systems. We illustrate their interpretation:

Denote by $P_{[a,b]}$, the $\perp$ projection onto the eigenvectors of H with eigenvalues $\lambda_i \in [a, b]$. Then,

$$\mathrm{Tr}(\rho_\beta P_{[a,b]}) = \mathbb{P}(\text{ system has energy, } E \in [a, b]) \tag{14}$$

i.e., upon observation, the energy value is measured to be in the interval $[a, b]$. More generally, observables in quantum mechanics are represented by bounded, and sometimes unbounded operators. We will usually assume that our observables come from $\mathcal{B}(\mathcal{H})$. This is the object that generalizes a random variable in classical probability and functions on phase space for a system of classical particles, or functions $f \in C(\Omega)$ for a classical spin system such as the Ising model. The mean of the observable A in the canonical Gibbs state at inverse temperature β is given by

$$\omega_\beta(A) = \mathrm{Tr}(\rho_\beta A) \tag{15}$$

to be compared with

$$\langle f \rangle_\beta = \int f d\mu_\beta \tag{16}$$

in the classical case. Similarly, the variance of A is given by

$$\omega_\beta([A - \omega_\beta(A)]^2) = \omega_\beta(A^2) - \omega_\beta(A)^2 \tag{17}$$

and the correlation (covariance) between A and B is given by

$$\omega_\beta(AB) - \omega_\beta(A)\omega_\beta(B) \tag{18}$$

In general, $A \in \mathcal{A} \subseteq \mathcal{B}(\mathcal{H})$, a norm-closed, unital *-subalgebra of $\mathcal{B}(\mathcal{H})$, a.k.a., a C^*-algebra. A map $\omega : \mathcal{A} \to \mathbb{C}$ is called a *state* if the system (a state on $\mathcal{A}$) if it is linear, positive, and normalized:

1. $\omega(A + cB) = \omega(A) + c\omega(B)$ for all $c \in \mathbb{C}$, $A, B \in \mathcal{A}$.
2. $\omega(A^*A) \geq 0$ for all $A \in \mathcal{A}$.
3. $\omega(\mathbb{1}) = 1$.

Linear functionals with the positivity and normalized properties above have further nice properties:

1. ω is continuous. i.e., ω is a bounded linear functional with
$$\|\omega\| = \sup_{\substack{\|A\|=1 \\ A \in \mathcal{A}}} |\omega(A)| = 1 \tag{19}$$
2. The Cauchy-Schwarz inequality holds:
$$|\omega(A^*B)| \leq \sqrt{\omega(A^*A)\omega(B^*B)} \tag{20}$$
3. $|\omega(A^*BA) \leq \omega(A^*A)\|B\|$.

3 The Free Energy Functional and the Variational Principle

Entropy plays an essential role in Thermodynamics and Statistical Mechanics as well as in Information Theory and in Large Deviation Theory in probability. It was Boltzmann who realized the precise connection between thermodynamic entropy and a mathematical formula involving physical states. Von Neumann introduced quantum entropy,

$$S(\rho) = -\operatorname{Tr}\rho\log\rho \tag{21}$$

for a density matrix ρ. One may observe that

$$0 \leq S(\rho) \leq +\infty \tag{22}$$

There is much to say about this function, and one could devote an entire course to its properties and applications. Here, we will use it to give another interpretation of the canonical formalism. Let's consider a finite quantum system with Hamiltonian H and define

$$F_\beta(\rho) = E(\rho) - \frac{1}{\beta}S(\rho) \tag{23}$$

where $E(\rho) = \operatorname{Tr}\rho H$. F_β is the free energy functional. At this point, it is not clear what it has to do with the previously defined quantity $\frac{1}{\beta}\log Z_\beta$, but here it is:

$$F_\beta(\rho_\beta) = \inf_\rho F_\beta(\rho) = -\frac{1}{\beta}\log\operatorname{Tr}e^{-\beta H} \tag{24}$$

with

$$\begin{aligned}\rho_\beta &= \frac{1}{Z}e^{-\beta H}\\ Z_\beta &= \operatorname{Tr}e^{-\beta H}\end{aligned}$$

Minimizing the free energy functional is equivalent to maximizing the entropy given the expectation of the energy. For a proof of (24) we need Klein's inequality.

Lemma 3.1 (Klein's Inequality)**.** *Let A and B be two non-negative definite matrices satisfying* $0 \leq A, B \leq \mathbb{1}$ *and such that* $\ker B \subset \ker A$*. Then*

$$\operatorname{Tr}A(\log A - \log B) \geq \operatorname{Tr}(A-B) + \frac{1}{2}\operatorname{Tr}(A-B)^2 \tag{25}$$

Proof. The function $f(x) = -x\log x$, $x > 0$, continuously extended such that $f(0) = 0$, is easily seen to be concave. In fact it is $C^2((0,\infty))$ with

$$f''(x) = -\frac{1}{x} \tag{26}$$

By the Taylor Remainder Theorem and the expression for f'', it follows that for all x and y such that $0 \leq x < y \leq 1$, there exists a ξ such that $x \leq \xi \leq y$ and

$$f(y) - f(x) - (y - x)f'(y) = -\frac{1}{2}(x - y)^2 f''(\xi) \geq \frac{1}{2}(x - y)^2 \tag{27}$$

As A and B are non-negative definite, they are diagonalizable. Denote their eigenvalues by a_i and b_i, and the corresponding orthonormal eigenvectors by φ_i and ψ_i, respectively. From the assumptions it follows that $0 \leq a_i, b_i \leq 1$. Using the spectral decompositions of A and B, i.e.,

$$A = \sum_i a_i |\varphi_i\rangle\langle\varphi_i| \tag{28}$$

$$B = \sum_i b_i |\psi_i\rangle\langle\psi_i| \tag{29}$$

$$\sum_i |\varphi_i\rangle\langle\varphi_i| = \sum_i |\psi_i\rangle\langle\psi_i| = \mathbb{1} \tag{30}$$

we see that

$$\begin{aligned}
&\operatorname{Tr} A(\log A - \log B) - \operatorname{Tr}(A - B) - \frac{1}{2}\operatorname{Tr}(A - B)^2 \\
&= \sum_{ij} \operatorname{Tr} |\psi_i\rangle\langle\psi_i||\varphi_j\rangle\langle\varphi_j| \left[-f(A) + f(B) + (A - B)f'(B) - \frac{1}{2}(A^2 + B^2 - 2AB) \right] \\
&= \sum_{ij} \operatorname{Tr} |\psi_i\rangle\langle\psi_i||\varphi_j\rangle\langle\varphi_j| \left[-f(a_j) + f(b_i) + (a_j - b_i)f'(b_i) - \frac{1}{2}(a_j - b_i)^2 \right] \\
&\geq 0
\end{aligned}$$

where the last inequality follows from applying (27) term by term. □

Now to prove the variational principle, we can apply Lemma 3.1 with $A = \rho$, where ρ is an arbitrary density matrix, and $B = \rho_\beta$. Note than $\ker B = \{0\}$. This gives

$$\beta(F_\beta - F(\beta)) = \operatorname{Tr} \rho \log \rho - \operatorname{Tr} \rho \log \left(\frac{e^{-\beta H}}{Z(\beta)} \right) \tag{31}$$

$$\geq \frac{1}{2} \operatorname{Tr}(\rho - \rho_\beta)^2 \geq 0 \tag{32}$$

If the RHS vanishes, we have $\rho = \rho_\beta$. Hence the minimum of F_β is uniquely attained for $\rho = \rho_\beta$.

4 Quantum Spin Systems

Quantum spin systems are the simplest examples of non-trivial interacting quantum systems with many degrees of freedom. They are defined on a finite set

Λ; typically $\Lambda \subseteq \mathbb{Z}^d$. Usually we have the same kind of spin in each $x \in \Lambda$, but it is sometimes useful to consider a more general situation where for all $x \in \Lambda$ we have a finite-dimensional Hilbert space $\mathcal{H}_x$ of dimension n_x; and where $\mathcal{H}_x \equiv \mathbb{C}^{n_x}$ and $\mathcal{H}_\Lambda = \bigotimes_{x\in\Lambda} \mathcal{H}_x$. In statistical mechanics we are typically interested in sequences of finite systems indexed by a sequence of sets Λ. For finite Λ, the Hamiltonian is a Hermitian matrix H_Λ and the observables are represented by matrices, regarded as linear transformations of $\mathcal{H}_\Lambda$: $\mathcal{A}_\Lambda = \mathcal{B}(\mathcal{H}_\Lambda)$. If $\Lambda_1 \subset \Lambda_2$, $\mathcal{A}_{\Lambda_1} \subset \mathcal{A}_{\Lambda_2}$ by the embedding provided by

$$\mathcal{A}_{\Lambda_1} \ni A \mapsto A \otimes \mathbb{1} \in \mathcal{A}_{\Lambda_1} \otimes \mathcal{A}_{\Lambda_2 \setminus \Lambda_1} \equiv \mathcal{A}_{\Lambda_2} \tag{33}$$

Often we consider sequences of finite $\Lambda \subseteq \mathbb{Z}^d$, or another countable set Γ. A natural way to describe the Hamiltonian H_Λ in such a case is by considering a "potential" or "interaction" of the following form: $\Phi : \{$ finite subsets of $\Gamma\} \to \bigcup_{\Lambda\subset\Gamma} \mathcal{A}_\Lambda$. where the union is defined as an inductive limit, and for all X, $\Phi(X) \in \mathcal{A}_X$, $\Phi(X) = \Phi(X)^*$, and $H_\Lambda = \sum_{X\subset\Lambda} \Phi(X)$. Boundary terms could also be added separately. e.g., in the Heisenberg model,

$$n_x = 2 \text{ for all } x \in \mathbb{Z}^d = \Gamma \tag{34}$$

$$\Phi(X) = 0 \text{ unless } X = \{x, y\}, |x - y| = 1 \tag{35}$$

$$\Phi(\{x, y\}) = -J\vec{\sigma_x} \cdot \vec{\sigma_y} \tag{36}$$

In this case, observe $\mathcal{A}_X \equiv \mathcal{A}_{X+a}$ for all $a \in \mathbb{Z}^d$ and $\Phi(X+a) = \tau_a(\Phi(X))$ where τ_a is the translation isomorphism mapping $\mathcal{A}_X$ to $\mathcal{A}_{X+a}$. τ_a is an automorphism of $*$-algebras:

$$\tau_a(A_\lambda B) = \tau_a(A) + \lambda\tau_a(B) \tag{37}$$

$$\tau_a(AB) = \tau_a(A)\tau_a(B) \tag{38}$$

$$\tau_a(A^*) = \tau_a(A)^* \tag{39}$$

$$\tau_a(\mathbb{1}) = \mathbb{1} \tag{40}$$

We will encounter other important examples of automorphisms shortly. (They describe the dynamics as well as the symmetries of the system.) As in the classical case, the dynamics of the system, i.e., its time evolution, is determined by the Hamiltonian H:

$$i\frac{d}{dt}\psi_t = H\psi_t \tag{Schrödinger}$$

$$\frac{d}{dt}A_t = i[H, A] \text{ for } A \in \mathcal{A} \subseteq \mathcal{B}(\mathcal{H}) \tag{Heisenberg}$$

where for convenience, we have set $\hbar = 1$.

The two equations are equivalent, in the sense that in moth cases, the solutions can be expressed in the one-parameter group of unitaries $U_t \in \mathcal{B}(\mathcal{H})$ generated by H as follows:

$$U_t = e^{-itH} \tag{41}$$

and then,

$$\psi_t = U_t\psi_0 \ , \ A_t = U_t^* A_0 U_t \tag{42}$$

Note that the map $A \mapsto \alpha_t^{(\Lambda)}(A) = e^{itH_\Lambda} A e^{-itH_\Lambda}$ where $A \in \mathcal{B}(\mathcal{H}_\Lambda)$ is an automorphism of $A_\Lambda = \mathcal{B}(\mathcal{H}_*)$. For models such as the Heisenberg model it is reasonable to ask whether it can be extended to an automorphism of $\bigcup_\Lambda A_\Lambda$, which we will call $\mathcal{A}_{\text{loc}}$: the algebra of local observables. The answer is no, but almost yes. What we need to do is to complete $\mathcal{A}_{\text{loc}}$, in the sense of metric spaces, to obtain a C^* algebra $\mathcal{A}$, which we now call the algebra of quasi-local observables. It contains all norm-limits of Cauchy sequences of local observables. It can be shown that there exist a one-parameter group of automorphisms α_t on $\mathcal{A}$ such that for all $A \in \mathcal{A}_{\text{loc}}$

$$\lim_{\Lambda \nearrow \mathbb{Z}^d} \alpha_t^{(\Lambda)}(A) = \alpha_t \tag{43}$$

in the norm topology. One says that $\alpha_t^\Lambda \to \alpha_t$ strongly. Note that there is no good way to define a Hamiltonian in the limit

$$\lim_{\Lambda \nearrow \mathbb{Z}^d} H_\Lambda =? \tag{44}$$

On the other hand, it is straightforward to define

$$\delta(A) = \lim_{\Lambda \nearrow \mathbb{Z}^d} [H_\Lambda, A] \tag{45}$$

for all $A \in \mathcal{A}_{\text{loc}}$. This does not automatically imply that one can define $e^{it\delta}(A)$ since δ is an unbounded operator. This is one reason why we need to define $\mathcal{A}$ on the $*$ completion (or closure) of $\mathcal{A}_{\text{loc}}$.

5 Symmetries and Symmetry Breaking (in the Heisenberg Model)

In finite volume, we have a unique Gibbs state $\sim e^{-\beta H_\Lambda}$; boundary conditions can be used to modify the ground states, i.e., $\lim_{\beta\to\infty} \rho_\beta$. Just like in the Ising model, we will ask whether we can also obtain different limits as $\Lambda \nearrow \mathbb{Z}^d$. It is instructive to look at ground states first and to look at the role of symmetries. Since boundary conditions (b_Λ) affect symmetries, they are more obvious in infinite system objects such as Φ or δ, rather than in the local Hamiltonians themselves. Let's focus on the spin $\frac{1}{2}$ Heisenberg model:

$$\vec{\sigma}\cdot\vec{\sigma} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{46}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{47}$$

$$= 2\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{48}$$

$$= 2t - \mathbb{1} \tag{49}$$

where $t(u\otimes v) = v\otimes u$. Clearly $[U\otimes U, t] = 0$, and this implies that $\vec{\sigma}\cdot\vec{\sigma}$ is $SU(2)$ invariant. More specifically, for all $U \in SU(2)$,

$$\pi_U^{(\Lambda)} = \bigotimes_{x\in\Lambda} U \tag{50}$$

so that $\pi_U^{(\Lambda)}$ is a unitary representation of $SU(2)$ on $\mathcal{H}_\Lambda$ and it commutes with the Hamiltonian H_Λ:

$$[H_\Lambda, \pi_U^{(\Lambda)}] = 0 \tag{51}$$

for all $U \in SU(2)$. The corresponding adjoint representation on $\mathcal{A}_\Lambda = \mathcal{B}(\mathcal{H}_\Lambda)$ is given by

$$\rho_U^{(\Lambda)}(A) = \pi_{U^*}^{(\Lambda)} A \pi_U^{(\Lambda)} \tag{52}$$

for all $A \in \mathcal{A}_\Lambda$. $\rho_U^{(\Lambda)}$ is a representation of $SU(2)$ by $*$-automorphisms of $\mathcal{A}_\Lambda$, and it commutes with the dynamics:

$$\rho_U^{(\Lambda)} \circ \alpha_t^{(\Lambda)} = \alpha_t^{(\Lambda)} \circ \rho_U^{(\Lambda)} \tag{53}$$

for all $U \in SU(2)$ and for all $t \in \mathbb{R}$.

As a consequence of the $SU(2)$ invariance of the dynamics, we also get the invariance of the finite volume Gibbs state:

$$\omega_\beta^{(\Lambda)}(A) = \frac{\operatorname{Tr} e^{-\beta H_\Lambda} A}{\operatorname{Tr} e^{-\beta H_\Lambda}} \tag{54}$$

satisfies

$$\omega_\beta^{(\Lambda)} \circ \rho_U^{(\Lambda)} = \omega_\beta^{(\Lambda)} \tag{55}$$

A major question in the statistical mechanics of the Heisenberg model is again the question of spontaneous symmetry breaking. As was the case in the Ising model, the symmetry will be broken only in sufficiently high dimension and at sufficiently low temperatures (large β). This topic will occupy the next several lectures.

Statistical Mechanics, Math 266: Week 5 Notes

February 2, 2010

1 Quantum Spin Systems: Existence of the Dynamics and Lieb-Robinson Bounds

Consider a relatively general class of systems (to keep notations simple, not the most general one can handle with the same arguments).

$$\Lambda \subset \mathbb{Z}^d, \text{ for all } x \in \mathbb{Z}^d, \mathcal{H}_x \equiv \mathbb{C}^n \tag{1}$$

$$\mathcal{H}_\Lambda = \bigotimes_{x\in\Lambda} \mathcal{H}_x, \mathcal{A}_\Lambda = \mathcal{B}(\mathcal{H}_\Lambda), \tag{2}$$

$$\mathcal{A}_{\Lambda_1} \subset \mathcal{A}_{\Lambda_2}, \text{ if } \Lambda_1 \subseteq \Lambda_2 \tag{3}$$

$$\mathcal{A}_{\text{loc}} = \bigcup_\Lambda \mathcal{A}_\Lambda, \mathcal{A} = \overline{\mathcal{A}_{\text{loc}}}^{\|\cdot\|} \tag{4}$$

$$H_\Lambda = \sum_{X\subset\Lambda} \Phi(X) \tag{5}$$

$$\text{For all } X, \Phi(X) = \Phi(X)^* \in \mathcal{A}_X \tag{6}$$

$$\alpha_t^{(\Lambda)}(A) = e^{itH_\Lambda} A e^{-itH_\Lambda} \tag{7}$$

Claim 1.1. *Under suitable conditions on Φ, there exist α_t on $\mathcal{A}$, strongly continuous one-parameter group of automorphisms on $\mathcal{A}$, such that for all $A \in \mathcal{A}_{loc}$*

$$\lim_{\Lambda\nearrow\mathbb{Z}^d} \alpha_t^{(\Lambda)}(A) = \alpha_t(A) \tag{8}$$

This limit will be a problem if there are more and more terms in

$$S_x(\Lambda) = \{X \subseteq \Lambda | x \in X, \Phi(X) \neq 0\} \tag{9}$$

and these elements $\Phi(X)$ are not decreasing in size fast enough as X increases. The limit (8) typically should not be expected to be trivial.

$$\operatorname{supp} \alpha_t^{(\Lambda)}(A) = \Lambda \tag{10}$$

Very often, Φ is of finite range. This means that $\Phi(X) = 0$ if $\operatorname{diam} X > R$ for some given range R, and where $\operatorname{diam} X = \max\{d(x,y) | x, y \in X\}$ and

$d(x,y) = |x-y| = \sum_{i=1}^{d} |x_i - y_i|$ in $\mathbb{Z}^d$. (Other definitions are possible). But in some cases pair interactions (or even higher-order interactions) that act at long distance need to be considered. e.g., magnetic dipole-dipole interaction decays as $\frac{1}{r^6}$.

Define $F(r) = \frac{1}{(1+r)^{d+1}}$ for $r \geq 0$ and define a norm on the interactions Φ by

$$\|\Phi\| = \max_{x,y\in\mathbb{Z}^d} \sum_{\substack{X\subseteq\mathbb{Z}^d \\ x,y\in X}} \frac{\|\Phi(X)\|}{F(|x-y|)} \tag{11}$$

In particular,

$$\sum_{\substack{X \\ x,y\in X}} \|\Phi(X)\| \tag{12}$$

has to decay at least as $\frac{1}{(|x-y|+1)^{d+1}}$.

Remark 1.1.

$$\sum_{x\in\mathbb{Z}^d} F(|x|) = \sum_{x} \frac{1}{(|x|+1)^{d+1}} \leq C \tag{13}$$

$$\sum_{z\in\mathbb{Z}^d} F(|x-z|)F(|z-y|) \leq \tilde{C}_\mu F(|x-y|) \leq 2^{d+1} C F(|x-y|) \tag{14}$$

where $\tilde{C}_\mu = \sup_{x,y} \sum_{z\in\mathbb{Z}^d} \frac{F(|x-z|)F(|z-y|)}{F(|x-y|)}$

The proof of the remark is left as a homework assignment.

For all $\mu > 0$ we define

$$F_\mu(r) = e^{-\mu r} F(r) \tag{15}$$

and observe that the inequalities (13) and (14) are also satisfied when we replace F with F_μ. But $\|\Phi\|_\mu$ defined with F_μ instead of F is of course a stronger norm with increasing μ.

Theorem 1.1 (Lieb-Robinson Bound)**.** *Suppose $\mu > 0$ such that $\|\Phi\|_\mu < +\infty$. Then there exists numbers C and v such that for all Λ, for all X,$Y \subset \Lambda$, and for all $A \in \mathcal{A}_X$, $B \in \mathcal{A}_Y$, and for all $t \in \mathbb{R}$, we have*

$$\|[\alpha_t^{(\Lambda)}(A), B]\| \leq \frac{2\|A\|\|B\|}{\tilde{C}_\mu} \sum_{\substack{x\in X \\ y\in Y}} F_\mu(d(x,y)) e^{v|t|} \tag{16}$$

where

$$\sum_{\substack{x\in X \\ y\in Y}} F(d(x,y)) \leq \min(|X|,|Y|) C e^{-\mu d(X-Y)} \tag{17}$$

and $C = \|\Phi\|_\mu$, $v = 2\tilde{C}_\mu \|\Phi\|_\mu$

Proof. Let $f(t)$ be the quantity $[\alpha_t^{(\Lambda)}, B]$. Then

$$f'(t) = [i[H_\Lambda, \alpha_t^{(\Lambda)}(A)], B] \tag{18}$$

Define $S_\Lambda(X) = \{Z \subseteq \Lambda | Z \cup X \neq \emptyset, \Phi(Z) \neq 0\}$ so that

$$\begin{aligned}
f'(t) &= i[\sum_{Z \in S_\Lambda(X)} [\alpha_t^{(\Lambda)}(\Phi(X)), \alpha_t^{(\Lambda)}(A)], B] && (19)\\
&= i \sum_{Z \in S_\Lambda(X)} -[[\alpha_t^{(\Lambda)}(A), B], \alpha_t^{(\Lambda)}(\Phi(X))] - [[B, \alpha_t^{(\Lambda)}(\Phi(Z))], \alpha_t^{(\Lambda)}(A)] && (20)\\
&= i[\overbrace{\sum_{Z \in S_\Lambda(X)} \alpha_t^{(\Lambda)}(\Phi(Z))}^{\widetilde{H}(t)}, f(t)] - i \sum_{Z \in S_\Lambda(X)} [\alpha_t^{(\Lambda)}(A), [\alpha_t^{(\Lambda)}(\Phi(Z)), B]] && (21)
\end{aligned}$$

where (21) uses the Jacobi identity.

Lemma 1.1. *For all $t \in \mathbb{R}$, let $\widetilde{H}(t) = \widetilde{H}(t)^* \in \mathcal{B}(\mathcal{H}_\Lambda)$ and $w(t) \in \mathcal{B}(\mathcal{H}_\Lambda)$. Then the unique solution $f(t)$ with initial condition $f(0)$ of*

$$f'(t) = i[\widetilde{H}(t), f(t)] + w(t) \tag{22}$$

is given by

$$f(t) = \tilde{\alpha}_t(f(0) + \int_0^t \alpha_{-s}(w(s))ds) \tag{23}$$

where $\tilde{\alpha}_t$ is the map solving

$$g'(t) = i[\widetilde{H}(t), g(t)] \tag{24}$$

i.e., $g(t) = \tilde{\alpha}_t(g(0))$.

Using the lemma,

$$\begin{aligned}
\|f(t)\| = \|f(0)\| + \| \int_0^t \tilde{\alpha}_{-s}(w(s))ds\| &\leq \|f(0)\| + \int_0^t ds \|w(s)\| && (25)\\
&\leq \|[A, B]\| + 2\|A\| \sum_{Z \in S_\Lambda(X)} \int_0^t ds \|[\alpha_s^{(\Lambda)}(\Phi(Z)), B]\| && (26)
\end{aligned}$$

Now define

$$C_B(X, t) = \sup_{\substack{A \in \mathcal{A}_X \\ A \neq 0}} \frac{\|[\alpha_t^{(\Lambda)}(A), B]\|}{\|A\|} \tag{27}$$

It follows now that

$$\|f(t)\| \leq \|A\| C_B(X, t) \tag{28}$$

and

$$C_B(X,t) \leq C_B(X,0) + 2 \sum_{Z \in S_\Lambda(X)} \|\Phi(Z)\| \int_0^t ds C_B(Z,s) \tag{29}$$

This inequality can be iterated. Note that

$$C_B(Z',0) \leq 2\|B\|\delta_Z(Z') \text{ for } B \in \mathcal{A}_Z \tag{30}$$

where

$$\delta_Z(Z') = \begin{cases} 1 & \text{if } Z \cup Z' \neq \emptyset \\ 0 & \text{if } Z \cup Z' = \emptyset \end{cases} \tag{31}$$

$$C_B(X,t) \leq C_B(X,0) + 2 \sum_{Z_1 \in S_\Lambda(X)} \|\Phi(Z_1)\| \int_0^t ds_1 \left(C_B(Z_1,0) + 2 \sum_{Z_2 \in S_\Lambda(Z_1)} \|\Phi(Z_2)\| \int_0^{s_1} ds_2 C_B(Z_2,s_2) \right) \tag{32}$$

$$\leq 2\|B\|\delta_Y(X) + 2 \sum_{Z_1 \in S_\Lambda(X)} 2t\|B\|\delta_Y(Z_1)\|\Phi(Z_1)\| + \tag{33}$$

$$+ 2\|B\| \sum_{Z_1 \in S_\Lambda(X)} \|\Phi(Z_1)\| \sum_{Z_2 \in S_\Lambda(Z_1)} \|\Phi(Z_2)\|\delta_Y(Z_2) \int_0^t ds \int_0^s ds_2 \tag{34}$$

$$\leq 2\|B\| \sum_{n=0}^{\infty} \frac{(2t)^n}{n!} a_n \tag{35}$$

where $a_n = \sum_{Z_1 \in S_\Lambda(X)} \cdots \sum_{Z_n \in S_\Lambda(Z_{n-1})} \delta_Y(Z_n) \prod_{i=1}^n \|\Phi(Z_i)\|$. Now, we estimate these coefficients:

$$a_1 = \sum_{Z \in S_\Lambda(X)} \delta_Y(Z)\|\Phi(Z)\| \tag{36}$$

$$\leq \sum_{y \in Y} \sum_{\substack{Z \in S_\Lambda(X) \\ y \in Z}} \|\Phi(Z)\| \tag{37}$$

$$\leq \sum_{y \in Y} \sum_{x \in X} F_\mu(d(x,y)) \sum_{\substack{Z \\ x,y \in Z}} \frac{\|\Phi(Z)\|}{F_\mu(d(x,y))} \tag{38}$$

$$\leq \|\Phi\|_\mu \sum_{y \in Y} \sum_{x \in X} F_\mu(d(x,y)) \tag{39}$$

$$\tag{40}$$

$$
\begin{aligned}
a_2 &= \sum_{Z_1 \in S_\Lambda(X)} \sum_{Z_2 \in S_\Lambda(Z_1)} \delta_Y(Z_2) \|\Phi(Z_1)\| \|\Phi(Z_2)\| && (41) \\
&\leq \sum_{y \in Y} \sum_{Z_1 \in S_\Lambda(X)} \|\Phi(Z_1)\| \sum_{z_1 \in Z_1} \sum_{\substack{Z_2 \\ z_1, y \in Z_2}} \|\Phi(Z_2)\| && (42) \\
&\leq \|\Phi\|_\mu \sum_{y \in Y} \sum_{z \in \Lambda} F_\mu(d(z,y)) \sum_{x \in X} F_\mu(d(x,z)) \sum_{\substack{Z_1 \\ x, z \in Z_1}} \frac{\|\Phi(Z_1)\|}{F_\mu(d(x,z))} && (43) \\
&\leq \tilde{C}_\mu \|\Phi\|_\mu^2 \sum_{\substack{y \in Y \\ x \in X}} F_\mu(d(x,y)) && (44)
\end{aligned}
$$

Proceeding in this way, one finds that for all $n \geq 1$,

$$
a_n \leq \|\Phi\|_\mu^n C_\mu^{n-1} \sum_{\substack{x \in X \\ y \in Y}} F_\mu(d(x,y)) \qquad (45)
$$

and this implies that $v = 2C\|\Phi\|_\mu$. $\square$

Statistical Mechanics, Math 266: Week 6 Notes

February 9 and 11, 2010

1 The Existence of Infinite System Dynamics

Corollary 1.1. *For $X \subset \Lambda$, define for $\delta > 0$*

$$X(t,\delta) = \{x \in \Lambda | d(x,X) \leq v|t| + \delta\}$$

Then there exists C such that for all $A \in \mathcal{A}_X$ and for all $t \in \mathbb{R}$, there exists $A_t(\delta) \in \mathcal{A}_{X(t,\delta)}$ such that

$$\|\alpha_t^{(\Lambda)}(A) - A_t(\delta)\| \leq C\|A\|e^{-\mu\delta}$$

Proof. The proof follows from the Lieb-Robinson bound and the following lemma:

Lemma 1.1. *Let $X \subset \Lambda$ and define $X^c = \Lambda \setminus X$. If $A \in \mathcal{A}_\Lambda$ satisfies*

$$\|[A,B]\| \leq \varepsilon\|B\|$$

for all $B \in \mathcal{A}_{X^c}$. Then there exists $A_\varepsilon \in \mathcal{A}_X$ such that $\|A - A_\varepsilon\| \leq \varepsilon$.

□

Theorem 1.1 (Existence of Infinite System Dynamics)**.** *Under the conditions described above, which imply Lieb-Robinson bounds for $\alpha_t^{(\Lambda)}$ uniformly in Λ, there exists a strongly continuous one parameter group of automorphisms, α_t on $\mathcal{A} = \overline{\mathcal{A}_{loc}}$, such that for all $A \in \mathcal{A}_{loc}$,*

$$\lim_{\Lambda \nearrow \mathbb{Z}^d} \|\alpha_t^{(\Lambda)}(A) - \alpha_t(A)\| = 0$$

The convergence is uniform on compact sets in t.

Proof. 1. The first step in the proof is to establish that for any increasing, absorbing sequence of finite sets, $\Lambda_n \nearrow \mathbb{Z}^d$, for fixed A and t, $(\alpha_t^{(\Lambda_n)})_{n\geq 1}$ is a Cauchy sequence in $\mathcal{A}$. Let $\Lambda_n \supset \Lambda_m$, then

$$\begin{aligned}
\alpha_t^{(\Lambda_n)}(A) - \alpha_t^{(\Lambda_m)}(A) &= \alpha_s^{(\Lambda_n)}\alpha_{t-s}^{(\Lambda_m)}(A)\Big|_0^t \\
&= \int_0^t ds \frac{d}{ds}(\alpha_s^{(\Lambda_n)}\alpha_{t-s}^{(\Lambda_m)}(A) \\
&= \int_0^t ds \quad i\alpha_s^{(\Lambda_n)}([H_{\Lambda_n}, \alpha_{t-s}^{(\Lambda_m)}(A)]) - i\alpha_s^{(\Lambda_n)}(\alpha_{t-s}^{(\Lambda_m)}([H_{\Lambda_m}, A]))
\end{aligned}$$

Hence

$$
\begin{aligned}
\|\alpha_t^{(\Lambda_n)}(A) - \alpha_t^{(\Lambda_m)}(A)\| &\leq \int_0^t ds \|[H_{\Lambda_n} - H_{\Lambda_m}, \alpha_{t-s}^{(\Lambda_m)}(A)]\| \\
&\leq \int_0^t ds \sum_{\substack{Y \subset \Lambda_n \\ Y \cup (\Lambda_n \setminus \Lambda_m) \neq \emptyset}} \|[\Phi(Y), \alpha_s^{(\Lambda_m)}(A)]\| \\
&\leq C\|A\|e^{\mu v|s|} \sum_{Y \subset \Lambda_n} \sum_{\substack{x \in X \\ z \in Y}} \|\Phi(Y)\| F_\mu(d(x,z)) \\
&\leq C\|A\|e^{\mu v|s|} \sum_{y \in \Lambda_n \setminus \Lambda_m} \sum_{\substack{Y \\ y \in Y}} \sum_{x \in X} \sum_{z \in Y} \|\Phi(Y)\| F_\mu(d(x,z)) \\
&\leq C\|A\|e^{\mu v|s|} \|\Phi\|_\mu \sum_{y \in \Lambda_n \setminus \Lambda_m} \sum_{z \in \mathbb{Z}^d} \sum_{x \in X} F_\mu(d(x,z)) F_\mu(d(y,z)) \\
&\leq C\|A\|e^{\mu v|s|} \|\Phi\|_\mu \sum_{y \in \Lambda_n \setminus \Lambda_m} F_\mu(d(x,y)) \to 0
\end{aligned}
$$

since F_μ is summable over $\mathbb{Z}^d$.

2. The limit $\alpha_t(A)$ is independent of the chosen sequence of (Λ_n).

3. $\alpha_t(A)$ defines an automorphism of $\mathcal{A}$, α_t satisfies the group property, and $\alpha_t(A)$ is continuous in t for all A.

□

2 Positive Semigroups on $\mathcal{A}$

In this section, we will not use that the Hilbert space $\mathcal{H}$ is finite-dimensional. The results, and the proofs given, are valid for arbitrary Hilbert spaces. In fact, $\mathcal{A}$ can be replaced by an arbitrary C^*-algebra $\mathcal{A}$, i.e., a subalgebra of $\mathcal{A}$ that is closed under taking and adjoints, and closed in the operator norm topology. It is convenient to assume that $\mathbb{1} \in \mathcal{A}$. For concreteness, you may still want to think of $\mathcal{A}$ as $M_n(\mathbb{C})$, but the finite-dimensionality will not be used.

Let $X \in \mathcal{A}$. Define $L_X \in \mathcal{B}(\mathcal{A})$, by

$$
L_X(A) = X^*AX - \frac{1}{2}(X^*XA + AX^*X) \tag{1}
$$

Clearly, as $\|L_X(A)\| \leq 2\|X\|^2\|A\|$, L_X is a bounded linear transformation on the Banach space $\mathcal{A}$. Therefore, we can define

$$
\gamma_t(A) = e^{tL_X}(A) = \left[\mathrm{id} + L_X + \frac{1}{2!}(L_X)^2 + \cdots\right](A)
$$

The family of transformations, $(\gamma_t)_{t \geq 0}$, is a *semigroup* and L_X is called the *generator* of γ_t. If the semigroup is differentiable in the sense that

$$
\frac{d}{dt}\gamma_t(A)
$$

exists for all $A \in \mathcal{A}$, then the generator is the linear map L defined by the derivative of the semigroup in $t = 0$:

$$L(A) = \frac{d}{dt}\gamma_t(A)\big|_{t=0}$$

In fact, for semigroups of this kind one can show that continuity in $t = 0$ implies differentiability. We will show that, if the generator is of the form given in (1), then the semigroup has the following following properties: $\gamma_t(\mathbb{1}) = \mathbb{1}$, and $\gamma_t(A)$ is positive definite if A is. A map γ_t with this property is called a positive map. We will prove these properties below. From these properties it follows that, for all states ω, and all $t \geq 0$, there is a state ω_t given by

$$\omega_t(A) = \omega(\gamma_t(A))$$

Although γ_t is a well-defined bounded linear transformation on $\mathcal{A}$ for all $t \in \mathbb{R}$, the properties that make it useful only hold for $t \geq 0$. E.g., the norm of the transformation γ_t diverges as $t \to -\infty$. Even more importantly, its positivity only holds for $t \geq 0$. So, although we have curves ω_t in the space of the linear functionals defined for all $t \in \mathbb{R}$, we will only use $t \geq 0$, as ω_t may cease to be state for $t < 0$. In infinite-dimensional situations one often considers γ_t with an *unbounded* generator L, in which case e^{tL} is often not defined for $t < 0$.

Proposition 2.1. *Let $X \in \mathcal{A}$ and let $L : \mathcal{A} \to \mathcal{A}$, be defined by*

$$L(A) = X^*AX - \frac{1}{2}(X^*XA + AX^*X)$$

Then, the following properties hold:
(i) For all $A \in \mathcal{A}$, $L(A^*) = L(A)^*$, and $\gamma_t(A^*) = \gamma_t(A)^*$, for all $t \geq 0$.
(ii) For all $A \in \mathcal{A}$, $L(A^*A) \geq L(A^*)A + A^*L(A)$.
(iii) The semigroup γ_t with generator L is unit preserving *(also called* unital*), i.e.,*

$$\gamma_t(\mathbb{1}) = \mathbb{1}, \text{ for all } t \geq 0,$$

(iv) The semigroup γ_t with generator L is positive, and satisfies, for all $A \in \mathcal{A}$

$$\gamma_t(A^*A) \geq \gamma_t(A^*)\gamma_t(A) = \gamma_t(A)^*\gamma_t(A) \geq 0, \text{ for all } t \geq 0,$$

(v) For every $t \geq 0$, γ_t is a contraction*, i.e., for all $A \in \mathcal{A}$*

$$\|\gamma_t(A)\| \leq \|A\|$$

Proof. *(i)* This follows directly from the definition of L and the properties of *.

(ii) This follows from the easy-to-verify identity

$$L(A^*A) - L(A^*)A - A^*L(A) = (XA - AX)^*(XA - AX)$$

(iii) This follows immediately form $L(\mathbb{1}) = 0$.

(iv) We will first prove that $\gamma_t(A^*A) \geq 0$, for all $A \in \mathcal{A}$, and then use that result to get the stronger property claimed in (iv).

We start by considering

$$0 \leq (\mathrm{id} + tL)(A^*)(\mathrm{id} + tL)(A)$$

which follows form (i). After expanding the product and using (ii) one obtains

$$0 \leq A^*A + tL(A^*A) + t^2 L(A^*)L(A)$$

from which we immediately get

$$0 \leq (\mathrm{id} + tL)(A^*A) + t^2\|L\|^2\|A\|^2$$

By applying the last inequality to the positive operator $\|A\|^2 - A^*A$, and using $\| \, \|A^*A\| - A^*A\| \leq \|A^*A\|$, we find

$$0 \leq \|A\|^2 - (\mathrm{id} + tL)(A^*A) + t^2\|L\|^2\|A\|^2.$$

By combining the last two inequalities one sees that

$$-t^2\|L\|^2\|A\|^2 \leq (\mathrm{id} + tL)(A^*A) \leq (1 + t^2\|L\|^2)\|A\|^2 \tag{2}$$

and *a fortiori*

$$-(1 + t^2\|L\|^2)\|A\|^2 \leq (\mathrm{id} + tL)(A^*A) \leq (1 + t^2\|L\|^2)\|A\|^2$$

from which it follows that

$$\|(\mathrm{id} + tL)(A^*A)\| \leq (1 + t^2\|L\|^2)\|A\|^2 \tag{3}$$

To prove the positivity of γ_t, we start from the expression

$$\gamma_t(A^*A) = \lim_{n\to\infty}(\mathrm{id} + \frac{t}{n}L)^n(A^*A)$$

Let $M(n) = 1 + t^2\|L\|^2/n^2$. By using (2) n times we get the following estimates:

$$\begin{aligned}(\mathrm{id} + \frac{t}{n}L)^n(A^*A) &\geq -(t^2/n^2)\|L\|^2\left[1 + M(n) + M(n)^2 + \cdots + M(n)^{n-1}\right]\|A\|^2 \\ (\mathrm{id} + \frac{t}{n}L)^n(A^*A) &\leq M(n)^n\|A\|^2 \end{aligned} \tag{4}$$

From (3), it follows that

$$\|(id + tL)\| \leq 1 + t^2\|L\|^2 \tag{5}$$

Next we will consider powers of the form $(id + sL)^k$. (5) gives

$$\|(id + sL)^k\| \leq M(s)^k$$

with $M(s) = 1 + s^2\|L\|^2$ and 2 gives

$$(id + sL)(A^*A) \geq -s^2\|L^2\|\|A\|^2$$

Claim 2.1. $(id{+}sL)^k(A^*A) \geq -s^2\|L\|^2\|A\|^2[1{+}M(s){+}M(s)^2{+}\cdots{+}M(s)^{k-1}]$

Proof. It holds for $k = 1$. Assume that it holds up to $k-1$; then

$$(id + sL)(A^*A) + s^2\|L\|^2\|A\|^2 \sum_{l=0}^{k-2} M(s)^l \geq 0$$

where the left hand side is equal to B^*B if you wish. Therefore,

$$(id + sL)(B^*B) = (id + sL)^k(A^*A) \geq -s^2\|L\|^2\|B^*B\|$$

and

$$\|B^*B\| = \|(id + sL)^{k-1}(A^*A) + s^2(\|L\|^2\|A\|^2 \sum_{l=0}^{k-2} M(s)^l\|$$

$$\leq M(s)^{k-1}\|A\|^2 + s^2\|L\|\|A\|^2 \sum_{l=0}^{k-2} M(s)^l$$

and $s = \frac{t}{n}$; Eventually $s^2\|L\|^2 \leq 1$, for such s

$$\leq s\|A\|^2 \sum_{l=0}^{k-1} M(s)^l$$

and

$$(id + sL)^k(A^*A) \geq -s^2\|L\|^2\|A\|^2 \sum_{l=0}^{k-1} M(s)^l$$

□

Now choose an s such that, $M(s) \leq 2$, so

$$(id + \frac{t}{n}L)^n(A^*A) \geq -2n\|L\|^2\|A\|^2\frac{t^2}{n^2} \to 0$$

Where by taking the limit $n \to \infty$ we see that γ_t , for $t \geq 0$, is a positive map.

Now, we will use the positivity to prove (iv) as follows. For any $A \in \mathcal{A}$, define,

$$f(t) = \gamma_t(A^*A) - \gamma_t(A)^*\gamma_t(A)$$

We need to show that $f(t) \geq 0$, for $t \geq 0$. As $f(0) = 0$, we have

$$f(t) = f(t) - \gamma_t(f(0)) = \int_0^t \frac{d}{ds}\left(\gamma_{t-s}(f(s))\right)\, ds$$

The compute the derivative in the integrand we use

$$\begin{aligned} \frac{d}{ds}\left(\gamma_{t-s}(f(s))\right) &= -\gamma_{t-s}(L(f(s))) + \gamma_{t-s}\frac{d}{ds}f(s) \\ \frac{d}{ds}f(s) &= L\gamma_s(A^*A) - L(\gamma_s(A))^*\gamma_s(A) - \gamma_s(A)^*L(\gamma_s(A) \end{aligned}$$

By using these relations we can write the integral as follows:

$$\int_0^t \frac{d}{ds}\left(\gamma_{t-s}(f(s))\right)\, ds = \int_0^t \gamma_{t-s}\left[L(\gamma_s(A)^*\gamma_s(A)) - L(\gamma_s(A)^*)\gamma_s(A) - \gamma_s(A)^* L(\gamma_s(A))\right]\, ds$$

By (ii) we know that the argument of γ_{t-s} is positive, and we already proved that γ_u is a positive map for $u \geq 0$. Hence, the integrand is positive for all $s \in [0, t]$, and it follows that $f(t) \geq 0$.

(v). By Lemma 2.1, proved below, we have

$$\|\gamma_t\| = \sup_{0 \neq A \in \mathcal{A}} \frac{\|\gamma_t(A^*A)\|}{\|A^*A\|}$$

We can use the norm inequality (4) to estimate the RHS as follows

$$\|(\mathrm{id} + \frac{t}{n}L)^n\| \leq (1 + \frac{t^2}{n^2}\|L\|^2)^n$$

By taking the limit $n \to \infty$ we obtain

$$\|\gamma_t\| \leq 1$$

□

Lemma 2.1. *Let T be a linear transformation on $\mathcal{A}$, satisfying $T(\mathbb{1}) = \mathbb{1}$, $T(A^*) = T(A)^*$, and $T(A^*A) \geq T(A^*)T(A)$, for all $A \in \mathcal{A}$. Then*

$$\|T\| = \sup_{0 \neq A \in \mathcal{A}} \frac{\|T(A^*A)\|}{\|A\|^2}$$

Proof. From the definition of $\|T\|$ as the supremum of $\|T(A)\|/\|A\|$, it follows that there is a sequence $A_n \in \mathcal{A}$, $\|A_n\| = 1$, such that $\|T\| = \lim_n \|T(A_n)\|$. For positive definite $A, B \in \mathcal{A}$, such that $B \leq A$, one has $\|B\| \leq \|A\|$. Using these properties we obtain

$$\limsup_n \|T(A_n^*A_n)\| \geq \limsup_n \|T(A_n^*)T(A_n)\| \geq \limsup_n \|T(A_n)^*\|\,\|T(A_n)\| = \|T\|^2$$

As $T(\mathbb{1}) = \mathbb{1}$, we must have $\|T\| \geq 1$, and therefore $\|T\|^2 \geq \|T\|$. In combination with the previous estimate this gives, and the fact that $\|A_n\| = 1$,

$$\sup_{0 \neq A \in \mathcal{A}} \frac{\|T(A^*A)\|}{\|A\|^2} \geq \limsup_n \|T(A_n^*A_n)\| \geq \|T\|.$$

The opposite inequality follows from the definition of $\|T\|$. □

3 Complete Positivity

This section is a brief aside on completely positive maps. A map $\gamma : \mathcal{A} \to \mathcal{A}$ is completely positive if

$$\gamma \otimes id_n : \mathcal{A} \otimes M_n \to \mathcal{A} \otimes M_n$$

is positive for all n. This is not a trivial definition, since there are positive maps which are not completely positive. As a concrete example, one may show that $\gamma(A) = A^t$ is positive, but $\gamma \otimes id_2$ is not positive. See Nielsen and Chuang, *Quantum Computation and Quantum Information* for more details.

Example 3.1. The γ_t defined in the previous section are all completely positive.

Example 3.2. Let $\alpha_t(A) = e^{itH} A e^{-itH}$ be the reversible dynamics of a system of the form $\mathcal{A} \otimes M_n$. Fix t and define $\gamma_t(A) = \mathrm{Tr}_{\mathbb{C}^n}\, \alpha_t(A) = \frac{1}{n} \sum_{i=1}^{n} id \otimes \varphi_i(\alpha_t(A))$, where $\varphi_i(\cdot) = \langle \varphi_i, \cdot \varphi_i \rangle$ and $\varphi_1, \ldots, \varphi_n$ form a basis for $\mathbb{C}^n$. In addition to complete positivity, we also have

$$\omega(\alpha_t(A \otimes \mathbb{1})) = \omega_1(\gamma_t(A \otimes \mathbb{1}))$$

Statistical Mechanics, Math 266: Week 7 Notes

February 18, 2010

1 Energy–entropy balance (EEB) inequalities

The characterization of thermal equilibrium that we will derive in this section is closely related to the variational principle. However, it will have the advantage that it can be formulated for infinite systems, while the variational principle suffers form the problem that the free energy functional diverges in the thermodynamic limit, so that it cannot be used, at least not without modification.

We will use the Energy–Entropy Balance inequalities to prive the Mermin–Wagner Theorem about absence of continuous symmetry breaking at strictly positive temperatures in dimensions ≤ 2 under quite general conditions.

The formulation of the EEB inequalities uses the function $f : [0, +\infty) \times [0, +\infty) \to (-\infty, +\infty]$ defined by

$$f(x,y) = \begin{cases} x \log \frac{x}{y} & \text{if } x, y > 0 \\ 0 & if x = 0, y \geq 0 \\ +\infty & if x > 0, y = 0 \end{cases} \tag{1}$$

In the following, whenever we write something of the form $x \log(x/y)$, we mean f as defined above. We will use the following elementary properties of f.

Proposition 1.1. *The function f defined in (??) has the following properties:*
(i) f is lower semicontinuous.
(ii) f is jointly convex in (x, y). i.e., for $\sum_i \alpha_i = 1$, $\alpha_i \geq 0$,

$$f(\sum_i \alpha_i x_i, \sum_i \alpha_i y_i) = \sum_i \alpha_i f(x_i, y_i)$$

(iii) f is homogeneous of degree one. i.e.,

$$f(\lambda x, \lambda y) = \lambda f(x, y)$$

(iv) For all finite sequences t_i, x_i, y_i, $i = 1, \ldots, n$, one has

$$f(\sum_i t_i x_i, \sum_i t_i y_i) \leq \sum_i t_i f(x_i, y_i)$$

We will formulate the EEB inequalities for a quantum system with Hilbert space $\mathcal{H}$, algebra of observables $\mathcal{A} = \mathcal{B}(\mathcal{H})$, and Hamiltonian H. The following theorem will be proved in the case $\dim \mathcal{H} < +\infty$, but is valid in a considerably more general setting.

Theorem 1.2. *Let ω be a state on $\mathcal{A}$. The following are equivalent conditions:*
(i) ω is the Gibbs state corresponding to H and inverse temperature β.
(ii) For all $X \in \mathcal{A}$ one has

$$\beta\omega(X^*[H,X]) \geq \omega(X^*X)\log\frac{\omega(X^*X)}{\omega(XX^*)} = f(\omega(X^*X),\omega(XX^*)) \tag{2}$$

Another way of stating the theorem is to say that the Gibbs state satisfies and is the only one that satisfies the inqualities (**??**) for all $X \in \mathcal{A}$. We will derive this property from the variational principle following a rather common procedure: we will define suitable curves in the space of all states that pass through the Gibbs state and compute and estimate the derivative of the free energy functional restricted to these curves. The EEB inequalities will follow from expressing that the state ω minimizes the free energy functional. The converse direction will be by explcit computation.

In order to define curves in the space of all states we recall the class of semigroups on $\mathcal{A}$ described in the previous lecture.

Let $X \in \mathcal{B}(\mathcal{H})$. Define $L_X : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, by

$$L_X(A) = X^*AX - \frac{1}{2}(X^*XA + AX^*X)$$

Clearly, as $\|L_X(A)\| \leq 2\|X\|^2\|A\|$, L_X is a bounded linear transformation on the Banach space $\mathcal{B}(H)$. Therefore, we can define

$$\gamma_t(A) = e^{tL_X}(A)$$

$(\gamma_t)_{t\geq 0}$ is a semigroup with the following properties: $\gamma(t)(\mathbb{1}) = \mathbb{1}$, and $\gamma_t(A)$ is positive definite if A is. A map γ_t with this property is called a positive map. From these properties it follows that, for all t, there is a unique density matrix ρ_t such that

$$\mathrm{Tr}\rho_t A = \mathrm{Tr}\rho\gamma_t(A)$$

In the finite-dimensional context, γ_t is a well-defined bounded linear transformation on $\mathcal{A}$ for all $t \in \mathbb{R}$. The norm of it, however, diverges as $t \to -\infty$. So although we have curves ρ_t, in the space of density matrices defined for all $t \in \mathbb{R}$, we will only use $t \geq 0$. In infinite-dimensional situations γ_t is in general not defined for $t < 0$.

Proof. The proof of the EEB inequalities consists in deriving the following two relations:

$$\lim_{t\downarrow 0}\frac{Tr\rho_t H - \mathrm{Tr}\rho_\beta H}{t} = \omega(X^*[H,X]) \tag{3}$$

$$\lim_{t\downarrow 0}\frac{S(\rho_t) - S(\rho_\beta)}{t} \geq \omega(X^*X)\log\frac{\omega(X^*X)}{\omega(XX^*)} \tag{4}$$

Here, $\rho_\beta = \rho_0$, and $\omega(A) = \mathrm{Tr}\rho A$. The EEB inequalities then follow from the VP:

$$F_\beta(\rho_t) - F_\beta(\rho_\beta) \geq 0$$

and therefore, for all $t > 0$, we must have

$$\frac{Tr\rho_t H - \mathrm{Tr}\rho_\beta H}{t} \geq \frac{1}{\beta}\frac{S(\rho_t) - S(\rho_\beta)}{t}$$

Assuming that the limits $t \downarrow 0$ exist, we get the EEB inequalities.

The derivative of the energy is easy to compute:

$$\frac{d}{dt}\omega(\gamma_t(H))\big|_{t=0} = \omega(L_X(H)) = \mathrm{Tr}\rho X^* H X - \frac{1}{2}\mathrm{Tr}\rho(X^*XH + HX^*X)$$

We are interested in the derivative in $\rho = \rho_\beta$. As $[\rho_\beta, H] = 0$, the last two terms are equal and can be combined. The result is (**??**).

For the entropy term we will need to differentiate operator valued functions of the type $\log A_t$. This is non-trivial. Usually the log function is defined by its series expansion around $\mathbb{1}$. To compute the derivative we will use the identity

$$\log x = \int_0^\infty \left[\frac{1}{1+t} - \frac{1}{x+t}\right] dt$$

for $x > 0$. So, for invertible $A_t \geq 0$, we consider

$$\begin{aligned}\frac{d}{dt}\log A_t &= \frac{d}{dt}\int_0^\infty \left[\frac{1}{1+s} - \frac{1}{A_t+s}\right] ds \\ &= \int_0^\infty (A_t+s)^{-1}\left(\frac{d}{dt}A_t\right)(A_t+s)^{-1} ds\end{aligned}$$

Here, we used the operator identity $A^{-1}(B-A)B^{-1} = A^{-1} - B^{-1}$ to compute

$$\frac{d}{dt}(A_t)^{-1} = -A_t^{-1}\left(\frac{d}{dt}A_t\right)A_t^{-1}$$

When we apply this to $-S(\rho_t)$ we get

$$\begin{aligned}\mathrm{Tr}\rho\frac{d}{dt}\log\rho_t\big|_{t=0} &= \mathrm{Tr}\rho\int_0^\infty \frac{1}{\rho+t}L_{X^*}(\rho)\frac{1}{\rho+t}dt \\ &= \mathrm{Tr}\rho\rho^{-1}L_{X^*}(\rho) \\ &= \mathrm{Tr}L_{X^*}(\rho)\end{aligned}$$

Now we can compute the derivative of the entropy term:

$$\begin{aligned}\frac{d}{dt}S(\rho_t)\big|_{t=0} &= -\mathrm{Tr}\frac{d}{dt}\rho_t\big|_{t=0} - \mathrm{Tr}\rho_t\frac{d}{dt}\log(\rho_t)\big|_{t=0} \\ &= -\mathrm{Tr}L_{X^*}(\rho)\log\rho - \mathrm{Tr}L_{X^*}(\rho) \\ &= -\mathrm{Tr}L_{X^*}(\rho)\log\rho\end{aligned}$$

where we used that $\mathrm{Tr} L_{X^*}(\rho) = \mathrm{Tr}\rho L_X(\mathbb{1}) = 0$.

Now we have to estimate (**??**). We will prove that

$$\begin{aligned} -\mathrm{Tr}\rho L_X(\log\rho) &= -\mathrm{Tr}\rho X^*(\log\rho)X + \frac{1}{2}\mathrm{Tr}\rho X^*X\log\rho + \frac{1}{2}\mathrm{Tr}\rho(\log\rho)X^*X \\ &\geq f(\mathrm{Tr}\rho X^*X, \mathrm{Tr}\rho XX^*) \end{aligned}$$

where f is the function defined in (**??**). To this end we use the spectral decomposition of ρ:

$$\rho = \sum_i \rho_i \, |\phi_i\rangle\langle\phi_i|$$

Using this we can write the LHS of the inequality as follows:

$$-\sum_{ij} \rho_i \, \langle \phi_i \mid X^*\phi_j\rangle \log\rho_j \, \langle\phi_j \mid X\phi_i\rangle + \sum_{ij} \rho_i \log\rho_i \, \langle\phi_i \mid X^*\phi_j\rangle \, \langle\phi_j \mid X\phi_i\rangle$$

If we let a_{ij} denote the matrix elements $\langle\phi_j \mid X\phi_i\rangle$, this can be written as

$$\sum_{ij} f(\rho_i, \rho_j)|a_{ij}|^2$$

Property (iv) of Proposition **??** then yields

$$\begin{aligned} -\mathrm{Tr}\rho L_X(\log\rho) &\geq f(\sum_{ij}\rho_i|a_{ij}|^2, \sum_{ij}\rho_j|a_{ij}|^2) \\ &= f(\mathrm{Tr}\rho X^*X, \mathrm{Tr}\rho XX^*) \end{aligned}$$

This concludes the proof of (i) $\Rightarrow$ (ii) in Theorem **??**. The opposite direction proceeds by solving the EEB inequalities. Suppose the Hamiltonian has eigenvalues λ_i and an orthonormal basis of eigenvectors ϕ_i. We will use the basis E_{ij} for the matrices:

$$E_{ij} = |\phi_i\rangle\langle\phi_j|\,, \quad E_{ij}* = E_{ji}, \quad E_{ij}E_{kl} = \delta_{jk}E_{il} \quad .$$

The spectral decomposition of the Hamiltonian can then be written as

$$H = \sum_i \lambda_i E_{ii}$$

First, we note that if ω satisfies (**??**), then the corresponding density matrix commutes with the Hamiltonian. This follows from the fact that the inequalities imply that, for all X,

$$\mathrm{Tr}\rho X^*HX - \mathrm{Tr}\rho X^*XH \in \mathbb{R}$$

and, as

$$\mathrm{Im}\ \mathrm{Tr}\rho X^*HX - \mathrm{Tr}\rho X^*XH = \mathrm{Tr}X^*X[\rho, H]$$

for arbitrary $X \in \mathcal{A}$, this implies $[\rho, H] = 0$. Hence, ρ has a spectral decomposition of the form

$$\rho = \sum_i \rho_i E_{ii}$$

Now, take $X = E_{ij}$ in the EEB inequalities. Then $[H, X] = (\lambda_i - \lambda_j)E_{ij}$, and the EEB inequality becomes:

$$\beta(\lambda_i - \lambda_j)\mathrm{Tr}\rho E_{jj} \geq F(\mathrm{Tr}\rho E_{jj}, \mathrm{Tr}\rho E_{ii})$$

By caculating the expecttations this is

$$\beta(\lambda_i - \lambda_j)\rho_j \geq F(\rho_j, \rho_i)$$

If $\rho_j \neq 0$, divide by it, and use the defintion of F to obtain:

$$\beta(\lambda_i - \lambda_j) \geq \log \frac{\rho_j}{\rho_i}$$

By combing this inequality with the one with the roles of i and j interchanged, we get, for all $i, j, \rho_i \neq 0, \rho_j \neq 0$,

$$\beta(\lambda_i - \lambda_j) = \log \frac{\rho_j}{\rho_i}$$

or, equivalently

$$\rho_i = \text{constant} \times e^{-\beta\lambda_i}$$

This completes the proof that ρ_β is the only density matrix satisfying the EEB inequalities for a fixed H and $\beta \geq 0$. $\square$

Statistical Mechanics, Math 266: Week 8 (Part 1) Notes

February 23, 2010

1 The Mermin-Wagner Theorem

We recall a few preliminaries.

Definition 1.1 (Kubo-Martin-Schwinger (KMS) condition). *ω satisfies the EEB if and only if $\omega(A\alpha_{i\beta}(B)) = \omega(BA)$ for all $A, B \in \mathcal{A}_{loc}$.*

The Gelfand-Naimark-Segal (GNS) representation is given as follows. Let ω be a state on a C^* algebra $\mathcal{A}$ (for example, an algebra of quasi-local observables). Then there exists a representation of $\mathcal{A}$, π_ω, on a Hilbert space $\mathcal{H}_\omega$ such that

$$\omega(A) = \langle \Omega_\omega, \pi_\omega(A)\Omega_\omega \rangle$$

where $\Omega_\omega \in \mathcal{H}_\omega$ is a cyclic vector for π_ω, i.e., $\pi_\omega(\mathcal{A})\Omega_\omega$ is dense in $\mathcal{H}_\omega$. Moreover, the representation with all these properties is unique up to a unitary transformation. If ω is α invariant, then α is implementable in π_ω: there exists a unitary operator on $\mathcal{H}_\omega$, U, such that

$$\pi_\omega(\alpha(A)) = U^* \pi_\omega(A) U$$

If α_t is strongly continuous and ω is α_t invariant, then there exists U_t which is also strongly continuous.

Theorem 1.2 (Stone-von Neumann). *If U_t is a strongly continuous group of unitaries in a Hilbert space $\mathcal{H}$, then there exists a densely defined self-adjoint operator H with domain $\mathrm{Dom}(H)$ such that*

$$U_t = e^{itH}$$

and $\psi \in \mathrm{Dom}(H)$ if

$$\lim_{t \searrow 0} \frac{U_t\psi - \psi}{t} = H\psi$$

exists.

Let H be a densely defined self-adjoint operator on a Hilbert space $\mathcal{H}$. Then there exists a resolution of the indentity E_λ such that

$$H = \int_{-\infty}^{+\infty} \lambda dE_\lambda$$

This is the generalization of the eigenvector decomposition of a compact operator. We also introduce the notation

$$P_{(a,b]} = \int_a^b dE_\lambda$$

which are orthogonal projections and $\mathbb{1} = \int_{-\infty}^{+\infty} dE_\lambda$.

Now, we continue towards the Mermin-Wagner Theorem. Let $\mathcal{A}$ be a C^* algebra such as the algebra of quasi-local observables of a quantum spin system on $\mathbb{Z}^d$, and suppose $\{\alpha_t\}_{t\in\mathbb{R}}$ is a strongly continuous one-parameter group of automorphisms of $\mathcal{A}$, which we will refer to as the *dynamics* of the system. The examples we have in mind are the dynamics of a quantum spin system generated by a not-too-long-range interaction Φ, e.g., one that satsifies, for some $\lambda > 0$,

$$\|\Phi\|_\lambda \equiv \sup_{x\in\mathbb{Z}^d} \sum_{x\ni X} e^{\lambda|X|}\|\Phi(X)\| \, < \, \infty \, .$$

A *symmetry* of the system is an automorphism, τ, of $\mathcal{A}$, which commutes with α_t, i.e.,

$$\alpha_t(\tau(A)) = \tau(\alpha_t(A)), \quad \text{for all } A \in \mathcal{A}, t \in \mathbb{R}$$

It is easy to see that if τ is a symmetry, than so is τ^{-1}. In fact, the set of all automorphisms commuting with the dynamics is a group for composition of automorphisms.

It is easy to see that if τ is a symmetry and ω is a $\beta-$KMS state for α_t, then $\omega \circ \tau$ is also $\beta-$KMS. The Mermin-Wagner-Hohenberg Theorem gives sufficient conditions that imply that all $\beta-$KMS states, ω, of the system are τ-invariant, i.e., $\omega(\tau(A)) = \omega(A)$, for all $A \in \mathcal{A}$. The original theorem, a special case of what we will prove here, says that no spontaneaous breaking of any continuous symmetry occurs at finite temperatures ($\beta < \infty$) in dimensions $d \leq 2$.

The general theorem involves the following two assumptions, which we will verify for a variety of systems, including two-dimensional models with a continuous symmetry.

MWH1: The symmetry τ is approximately inner in the sense that there exist a sequence of unitaries $U_n \in \mathcal{A}$ such that

$$\lim_{n\to\infty} \|\tau(A) - U_n^* A U_n\| = 0, \quad \text{for all } A \in \mathcal{A} \, .$$

We also assume that these unitaries can be taken form the domain of δ, the generator of the dynamics $\alpha_t = e^{it\delta}$. Equivalently, we assume that the follwing limits exist

$$\lim_{t\to 0} \frac{\alpha_t(U_n) - U_n}{t} = i\delta(U_n) \, .$$

Note that it follows from these assumptions that τ^{-1} is also approximately inner, approximated by the unitaries U_n^*, and that $U^* \in \mathrm{Dom}(\delta)$.

The second assumption comes in two versions.

MWH2: We assume that one of the following holds:
(i) there exists a constant M such that $\|\delta(U_n)\| \leq M$, for all n, or
(ii) all $\beta-$KMS states are τ^2-invariant and there exists a constant M such that

$$\|U_n^*\delta(U_n) + U_n\delta(U_n^*)\| \leq M, \quad \text{for all } n\,.$$

Theorem 1.3. *Suppose τ is a symmetry of the system $(\mathcal{A}, \alpha_t)$ such that conditions MWH1 and MWH2 ((i) or (ii)) are satisfied. Then, all $\beta-$KMS states are $\tau-$ invariant for all $\beta \in [0, \infty)$.*

Using the assumptions and the EEB inequalities, we will prove that if ω is $\beta-$KMS, then there exists a constant C such that

$$\omega \circ \tau(A^*A) \leq C\omega(A^*A) \tag{1}$$

The constant C will depend only on β and M. This a uniform version of absolute continuity of $\omega \circ \tau$ with respect to ω. It is not hard to prove that for extremal $\beta-$KMS states one has the dichotomy: either they are equal or they are disjoint. That is, if they are quasi-equivalent states, *a fortiori*, if one is absolutely continuous with respect to the other, then they are necessarily equal. This follows from the general result that (**??**) implies that there exists $0 \leq T \in \pi_\omega(\mathcal{A})' \cap \pi_\omega(\mathcal{A})''$ such that $\omega \circ \tau(A^*A) = \langle \Omega_\omega, \pi_\omega(A^*A)T\Omega_\omega\rangle$. Since extremal KMS states are factor states, such T must be a multiple of $\mathbb{1}$ and, therefore, $\omega \circ \tau = \omega$.

So, from (**??**), it will follow that all extremal $\beta-$KMS states are $\tau-$invariant and, therefore, by taking convex combinations, all $\beta-$KMS states are $\tau-$invariant.

The second version of MWH2 includes the assumption that we already know that the $\beta-$KMS states are τ^2-invariant. This is no restriction for compact continuous summetry groups. For discrete groups such as finite groups or lattice translations one needs version (i). In general, (i) implies (ii), but note that for involutions ($\tau^2 = \mathbb{1}$), (i) and (ii) are equivalent. Now, we prove Theorem **??**.

Proof. Let ω be a $\beta-$KMS state. To prove (**??**) we will use the EEB inequalities and the GNS reprensentation of ω. As any KMS state is time invariant, α_t is unitarily implemented by unitaries U_t in the GNS representation. As α_t is strongly continuous, U_t is a strongly continuous one-parameter group generated by a s.a. operator H, with dense domain $\mathrm{Dom}(H)$, and such that $H\Omega = 0$, where Ω is the cyclic vector representing ω. We will need the spectral resolution of H:

$$H = \int \lambda\, dE_\lambda$$

to define a resolution of the identity by mutually orthogonal projections $P_n, n \in \mathbb{Z}$, $\sum_n P_n = \mathbb{1}$, as follows

$$P_n = \int_{(n,n+1]} dE_\Lambda$$

It is clear that, to prove (**??**), it is sufficient to prove that there exists a constant C, independent of n, such that for all $A \in \mathcal{A}$

$$\omega \circ \tau(A^* P_n A) \leq C\omega(A^* P_n A)$$

or more accurately, we will prove that, for all m, n, and for $A \in \mathcal{A}_0$, a norm-dense $*-$subalgebra of $\mathcal{A}$, we have

$$\langle \Omega \mid U_m^* \pi(A^*) P_n \pi(A) U_m \Omega \rangle \leq C \langle \Omega \mid \pi(A^*) P_n \pi(A) \Omega \rangle \tag{2}$$

By summing over n and taking the limit $m \to \infty$ one obtains (**??**).

To prove (**??**) we need the following estimates for quantities that appear in the EEB inequalities. For convenience, we introduce the notation $A_n = P_n \pi(A)$. For the first estimate, note that vectors of the form $A_n \Omega$ are in the domain of H. We will also use $\omega(\,\cdot\,)$ as shorthand for $\langle \Omega \mid \,\cdot\, \Omega \rangle$. Then we have, by using $H\Omega = 0$,

$$\begin{aligned}
\omega(A_n^* \delta(A_n)) &= \langle \Omega \mid A_n^* P_n H P_n A_n \Omega \rangle \\
&\leq (n+1) \langle \Omega \mid A_n^* P_n H P_n A_n \Omega \rangle \\
&\leq (n+1)\, \omega(A_n^* A_n)
\end{aligned}$$

For the entropy term, we first observe that, using the KMS condition, we can relate $\omega(A_n^* A_n)$ and $\omega(A_n A_n^*)$ as follows:

$$\begin{aligned}
\omega(A_n A_n^*) &= \omega(A_n^* \alpha_{i\beta}(A_n)) \\
&= \langle \Omega \mid A_n^* P_n E^{-\beta H} P_n A_n \Omega \rangle \\
&\leq e^{-\beta n} \omega(A_n^* A_n)
\end{aligned}$$

From this estimate we get

$$\begin{aligned}
\omega(A_n^* A_n) \log \frac{\omega(A_n^* A_n)}{\omega(A_n A_n^*)} &\geq \omega(A_n^* A_n) \log \frac{\omega(A_n^* A_n)}{e^{-\beta n} \omega(A_n^* A_n)} \\
&\geq \beta n\, \omega(A_n^* A_n)
\end{aligned}$$

The EEB inequality for the observable $X = U_m A_n$:

$$\beta \omega(A_n^* U_m^* \delta(U_m A_n)) \geq \omega(A_n^* A_n) \log \frac{\omega(A_n^* A_n)}{\omega(U_m A_n A_n^* U_m^*)}$$

By using the derivation property on the left and adding and substracting a term on the right, and reorganizing this can be written as (watch the stars!)

$$\omega(A_n^* A_n) \log \frac{\omega(A_n^* A_n)}{\omega(U_m A_n A_n^* U_m^*)} \leq \beta \omega(A_n^* U_m^* \delta(U_m) A_n) + \beta \omega(A_n^* \delta(A_n)) - \omega(A_n^* A_n) \log \frac{\omega(A_n^* A_n)}{\omega(A_n^* A_n)}$$

The last two terms can be bounded by the estimates we prepared. The result gives

$$\omega(A_n^* A_n) \log \frac{\omega(A_n A_n^*)}{\omega(U_m A_n A_n^* U_m^*)} \leq \beta \omega(A_n^* U_m^* \delta(U_m) A_n) + \beta \omega(A_n^* A_n) \tag{3}$$

Now it is time to use MWH2. The two versions are treated slightly differently. With version (i), we immediately get

$$\omega(A_n^* A_n) \log \frac{\omega(A_n A_n^*)}{\omega(U_m A_n A_n^* U_m^*)} \leq \beta(M+1)\omega(A_n^* A_n).$$

After simplifying, exponentiating, and reversing the roles of A_n and A_n^*, as well as τ and τ^{-1}, one gets (**??**) with $C = e^{\beta(M+1)}$.

In order to use MWH2 (ii), we use (**??**) and the similar bound for U_m and U_m^* interchanged. By adding the two bounds we get:

$$\begin{aligned} & \omega(A_n^* A_n) \log \frac{\omega(A_n A_n^*)^2}{\omega(U_m A_n A_n^* U_m^*)\omega(U_m^* A_n A_n^* U_m)} \\ \leq \quad & \beta\omega(A_n^*[U_m^*\delta(U_m) + U_m\delta(U_m^*)]A_n) + 2\beta\,\omega(A_n^* A_n) \end{aligned}$$

In the same way as before, but by using (ii) instead of (i), we obtain

$$\omega(A_n A_n^*)^2 \leq e^{\beta(M+2)}\omega(\tau(A_n A_n^*)\omega(\tau^{-1}(A_n A_n^*))$$

As $\omega \circ \tau$ is a $\beta-$KMS state, too, we can write

$$\omega(\tau(A_n A_n^*))^2 \leq e^{\beta(M+2)}\omega(\tau^2(A_n A_n^*)\omega(A_n A_n^*)$$

Now, we have to used that $\beta-$KMS states are τ^2-invariant. By taking square roots we get (**??**) with $C = e^{\beta(M+2)/2}$. □

2 Applications. The Mermin-Wagner-Hohenberg Theorem

Recall that the assumption MWH2 of Theorem **??** came in two versions. We assumed that one of the following holds: (i) there exists a constant M such that $\|\delta(U_n)\| \leq M$, for all n; (ii) all $\beta-$KMS states are τ^2-invariant and there exists a constant M such that

$$\|U_n^*\delta(U_n) + U_n\delta(U_n^*)\| \leq M, \quad \text{for all } n\,.$$

We still need to show how the theorem is used to prove absence of continuous symmetry breaking in two dimensions at finite temperature. As we will show a bit further, MWH2 (ii), but not (i), can be verified in this case. For an example of the first version of MWH2, see the homework. The following lemma allows us to apply the main theorem to continuous symmetries.

Lemma 2.1. *Let $\{tau_\phi \mid \phi \in S^1\}$ be a compact connected continuous one-parameter group of automorphisms of $\mathcal{A}$. Let K be a set of states ω such that $\omega \circ \tau_\phi^2 = \omega$ implies $\omega \circ \tau_\phi = \omega$, for any $\phi \in S^1$. Then all $\omega \in K$ are $\tau_\phi-$invariant for all $\phi \in S^1$.*

Proof. As $\tau_\pi^2 = \mathrm{id}$, the assumptions imply that $\omega \circ \tau_\pi = \omega$. By repeating the argument n more times we get that $\omega \circ \tau_{\pi/2^n}$. It follows immediately that ω is invariant for all τ_ϕ with ϕ of the form $\phi = \sum_{n=0}^{N} a_n 2^{-n}\pi$, where $a_n \in \mathbb{Z}$. Clealry, such ϕ form a dense set in S^1. Now, for every $A \in \mathcal{A}$, the function $\phi \to \omega(\tau_\phi(A) - A)$ is continous and vanishes on dense subset of S^1. Hence, it vanishes everywhere. □

For symmetries representing an arbitrary compact Lie group, we can apply this lemma for a generating set of one-dimensional compact subgroups.

Now, we will verify MWH1 and MWH2 (ii) for two-dimensional quantum spin systems with a connected compact continuous symmetry group. For simplicity we will consider pair interactions only. This means that the dynamics is generated by local Hamiltonians of the form

$$H_\Lambda = \sum_{x,y\in\Lambda} J(x,y)\Phi_{x,y} \tag{4}$$

for finite subsets Λ in $\mathbb{Z}^2$, where $\Phi_{x,y} \in \mathcal{A}_{\{x,y\}}$ are assumed to be uniformly bounded: say $\|\Phi_{x,y}\| \leq 1$, for all $x, y \in \mathbb{Z}^2$. An example of such a Hamiltonian is the Heisenberg model. Boundary terms are irrelevant in our considerations here. Suppose that there there are unitary representations

$$U_x(\phi) = e^{i\phi X_x}, \phi \in S^1,$$

with generators $X_x = X_x^* \in \mathcal{A}_{\{x\}}$, $x \in \mathbb{Z}^2$. E.g., for spin rotations the generators are SU(2) spin matrices.

Consider the boxes $\Lambda_m = [-m, m]^2 \subset \mathbb{Z}^2$. It is easy to satisfy MWH1 with a sequence of unitaries of the form

$$U_m(\phi) = \bigotimes_{x\in\Lambda_{2m}} U_x(\phi_m(x))$$

where $\phi_m(x) = \phi$, for all $x \in \Lambda_m$ and, for the moment, arbitrary for $x \in \Lambda_2 m \setminus \Lambda_m$.

Translation invariance is not required; in fact the argument works for inhomogeneous systems with spins of different magnitudes at different sites. We will assume that there is a uniform bound on the norm of the generators, say, there is a constant G such that $\|X_x\| \leq G$, for all $x \in \mathbb{Z}^2$.

Proposition 2.2. *For a quantum spin system on $\mathbb{Z}^2$ with local Hamiltonians of the form (**??**), with coupling constants $J(x,y)$, satisfying*

$$\sup_x \sum_{y\in\mathbb{Z}^2} |x-y|^2 |J(x,y)| < +\infty$$

we can find $\phi_m(x)$ such that there exists a constant M such that

$$\|U_m\delta(U_m^*) + U_m^*\delta(U_m)\| \leq M, \text{ for all } m$$

*and Theorem **??** can be applied.*

Proof. The idea behind the choice of the unitaries U_n that approximate the symmetry transformation is that, in the case of continuous symmetries such as a rotation by an angle ϕ, it is possible to interpolate "smoothly" between rotations by a fixed angle in any given finite volume, and zero rotation at infinity, in such a way that there is a uniform bound on the energy involved in such a perturbation.

Claim: it suffices to take ϕ_m defined as follows:

$$\phi_m(x) = \begin{cases} \phi & \text{if } x \in \Lambda_m \\ (2 - \frac{\min\{|x_1|,|x_2|\}}{m})\phi & \text{if } x \in \Lambda_{2m} \setminus \Lambda_m \\ 0 & \text{if } x \in x \notin \Lambda_{2m} \end{cases}$$

The quantity we need to bound is the following:

$$\|U_m \delta(U_m^*) + U_m^* \delta(U_m)\| \leq \sum_{x,y} |J(x,y)| \|\Delta_{x,y}\|$$

where

$$\begin{aligned} \Delta_{x,y} &= U_x(\phi_m(x)) U_y(\phi_m(y)) \Phi_{x,y} U_x(\phi_m(x))^* U_y(\phi_m(y))^* \\ &\quad + U_x(\phi_m(x))^* U_y(\phi_m(y))^* \Phi_{x,y} U_x(\phi_m(x)) U_y(\phi_m(y)) - 2\Phi_{x,y} \end{aligned}$$

By expanding the unitaries we can rewrite this as follows:

$$\begin{aligned} \Delta_{x,y} &= 0 + i[\phi_m(x) X_x + \phi_m(y) X_y, \Phi_{x,y}] - i[\phi_m(x) X_x + \phi_m(y) X_y, \Phi_{x,y}] \\ &\quad + 2 \sum_{n \geq 1} \frac{(-1)^n}{(2n)!} \mathrm{ad}^{2n}_{\phi_m(x) X_x + \phi_m(y) X_y}(\Phi_{x,y}) \end{aligned}$$

The trick is to realize that $\Delta_{x,y}$ only depends on the differences $\phi_m(x) - \phi_m(y)$. This can be seen as follows:

$$\phi_m(x) X_x + \phi_m(y) X_y = \frac{1}{2}(\phi_m(x) + \phi_m(y))(X_x + X_y) + \frac{1}{2}(\phi_m(x) - \phi_m(y))(X_x - X_y)$$

Let us call the first term $A_{x,y}$ and the second term $B_{x,y}$. Then, it is easily checked that $A_{x,y}$ and $B_{x,y}$ commute. Hence,

$$\mathrm{ad}_{\phi_m(x) X_x + \phi_m(y) X_y} = \mathrm{ad}_{A_{x,y}} + \mathrm{ad}_{B_{x,y}}$$

with $\mathrm{ad}_{A_{x,y}}$ and $\mathrm{ad}_{B_{x,y}}$ commuting as well. By assumption we have $\mathrm{ad}_{A_{x,y}}(\Phi_{x,y}) = 0$. Using these properties we can derive the following estimate for $\|\Delta_{x,y}\|$:

$$\|\Delta_{x,y}\| \leq 2 \sum_{n \geq 1} \frac{1}{(2n)!} \left(\frac{\phi_m(x) - \phi_m(y)}{2} \right)^{2n} \|\mathrm{ad}^{2n}_{B_{x,y}}(\Phi_{x,y})\| \tag{5}$$

Since $d = 2$, we have $|\Lambda_m| = (2m+1)^2$. Also not that $|\phi_m(x) - \phi_m(y)| \leq |\phi|/m$.

Therefore, the sum over x can be estimated by

$$\begin{aligned}
\|\Delta_{x,y}\| &\leq 2 \sum_{|x|\leq 2m, y\in\mathbb{Z}^2} |J(x,y)| \left(\frac{|x-y|}{2m}\right)^2 \sum_{n\geq 1} \frac{1}{(2n)!} (2|\phi| \|B_{x,y}\|)^{2n} \|\Phi_{x,y}\| \\
&\leq 4 \sum_{|x|\leq 2m, y\in\mathbb{Z}^2} |x-y|^2 |J(x,y)| \frac{\|B_{x,y}\|^2}{(2m)^2} e^{4|\phi| \|B_{x,y}\|^2} \\
&\leq \text{constant} \times \sup_x \sum_{y\in\mathbb{Z}^2} |x-y|^2 |J(x,y)|
\end{aligned}$$

□

One can obtain a similar condition on $J(x,y)$ that excludes continuous symmetry breaking in one dimension.

Statistical Mechanics, Math 266: Week 8 (Part 2) Notes

February 25, 2010

1 Ideal Gases

Ideal gases are point particles, non-interacting (no forces between them), only subject to being enclosed in a finite volume, i.e. boundary conditions. The particles only have kinetic energy. There is an important issue, however, related to how to compute the entropy, or in other terms, how to "count states". The parameters are continuous, so counting is not obvious in the classical case. To make this more clear, let's consider first the standard classical treatment of the classical ideal (or free) gas of point particles of mass m. Consider N particles in $\Lambda \subseteq \mathbb{R}^d$. Configurations are given by $(p,q) \in (\mathbb{R}^d)^{2N}$ where $p_1, \ldots, p_N \in \mathbb{R}^d$ and $q_1, \ldots, q_N \in \Lambda$ are the momenta and positions of particles $1, \ldots, N$. Recall that

$$H(\mathbf{p},\mathbf{q}) = \sum_{i=1}^{N} \frac{1}{2m} p_i^2 \tag{1}$$

$$\mu_\Lambda(\mathbf{p},\mathbf{q}) = \frac{1}{Z_\Lambda} e^{-\beta H(\mathbf{p},\mathbf{q})} = \frac{1}{Z} \prod_{i=1}^{N} (e^{-\frac{\beta p_i^2}{2m}}) \tag{2}$$

are independent of $\mathbf{q}$, and

$$Z(\beta) = \int_{\mathbb{R}^d} e^{-\frac{\beta p_i^2}{2m}} d\mathbf{p} = (\frac{2m\pi}{\beta})^{d/2} \tag{3}$$

The free energy and free energy density functions are

$$F(\Lambda, N, \beta) = -\frac{1}{\beta} \log Z_\Lambda = -\frac{1}{\beta} N \log |\Lambda| Z(\beta) \tag{4}$$

$$f(\rho,\beta) = \lim_{\substack{\Lambda \nearrow \mathbb{R}^d \\ N \to \infty \\ \rho = N/|\Lambda|}} \frac{F(\Lambda, N, \beta)}{|\Lambda|} = -\frac{1}{\beta} \rho \log Z(\beta) - \frac{1}{\beta} \rho \log |\Lambda| \tag{5}$$

Note that the free energy density does not converge in the thermodynamic limit. This is not what we expect, but why is it wrong? Before considering how to fix this point, let us focus, as Gibbs originally did, on the entropy term in the free energy. Our investigation will lead us to the Gibbs paradox.

$$F = E - \frac{1}{\beta} S \tag{6}$$

We need to calculate E in order to get S from F.

$$\begin{aligned} E &= \int H(\mathbf{p}, \mathbf{q}) \mu(d\mathbf{p}, d\mathbf{q}) &(7)\\ &= \frac{N \int_{\mathbb{R}^d} \frac{1}{2m} p^2 e^{-\frac{\beta p^2}{2m}} d^d p}{\int_{\mathbb{R}^d} e^{-\frac{\beta p^2}{2m}} d^d p} &(8)\\ &= \frac{N}{\beta} \frac{\int_{\mathbb{R}^d} x^2 e^{-x^2} d^d x}{\int_{\mathbb{R}^d} e^{-x^2} d^d x} &(9)\\ &= \frac{dN}{2\beta} = \frac{d}{2} N k_B T &(10) \end{aligned}$$

where we have made the substitution $x = \sqrt{\frac{\beta}{2m}} p$. Thus, we have

$$\begin{aligned} S(\Lambda, N, \beta) &= -\beta F + \beta E &(11)\\ &= N \log |\Lambda| Z(\beta) + \frac{d}{2} N &(12) \end{aligned}$$

The $\log |\Lambda|$ dependence is problematic because the "thermodynamic" entropy should be extensive; also, if we mix two identical ideal gases, their entropies should just add. Consider a system contained in two parts $\Lambda = \Lambda_1 \cup \Lambda_2$ where $\Lambda_1 \cap \Lambda_2 = \emptyset$ each with $N_i = \rho |\Lambda_i|$ particles and the same β. Letting $N = N_1 + N_2$,

$$S_{\text{tot}} = \begin{cases} N(\log Z(\beta) + \frac{d}{2}) + \rho|\Lambda_1| \log |\Lambda_1| + \rho |\Lambda_2| \log |\Lambda_2| & \text{(Same atoms)} \\ N(\log Z(\beta) + \frac{d}{2}) + N \log(|\Lambda_1| + |\Lambda_2|) & \text{(Different atoms)} \end{cases} \tag{13}$$

where the two distinct expressions for the entropy is Gibb's paradox. The interesting fact is that it is *correct* and experimentally verified in each of the cases indicated above. The difference between the entropy of the two cases,

$$\Delta S = N_1 \log \frac{|\Lambda|}{|\Lambda_1|} + N_2 \log \frac{|\Lambda|}{|\Lambda_2|} \tag{14}$$

is called the entropy of mixing. Gibbs proposed to circumvent these issues by the inclusion of a factor of $\frac{1}{N!}$ in the integral over phase space. Then the partition function becomes

$$Z_\Lambda \to \frac{1}{N!} Z_\Lambda$$

and the free energy becomes

$$\begin{aligned} F(\Lambda, N, \beta) &\to F(\Lambda, N, \beta) + \log \frac{1}{N!} \\ &\sim F - N \log N + N \end{aligned}$$

and with $N = \rho|\Lambda|$, the $|\Lambda|$ dependent term is cancelled. As for the entropy,

$$S(\Lambda, N, \beta) = N \log |\Lambda| Z(\beta) + \frac{d}{2} N - N \log N + N \tag{15}$$
$$= N \log \frac{|\Lambda|}{N} Z(\beta) + (\frac{d}{2} + 1) N \tag{16}$$
$$= \rho|\Lambda| \log \rho Z(\beta) + (\frac{d}{2} + 1) \rho |\Lambda| \tag{17}$$

it becomes proportional to $|\Lambda|$.

2 Identical Particles in Quantum Mechanics

In Quantum Mechanics, the starting point of our description is the Hilbert space of states. If you have two identical particles in the same box, the states should be

$$\psi \in \mathcal{H} \otimes \mathcal{H}$$

where $\mathcal{H}$ is the Hilbert space for one particle. "Needless to say (but disconcerting at times), identical particles cannot be distinguished", otherwise they would not be truly identical.

So what is the difference between $\psi = \varphi_1 \otimes \varphi_2$ and $\psi = \varphi_2 \otimes \varphi_1$? Indeed, all ψ obtained by permuting indices should be equivalent. It turns out that to define the space of particles, we have to restrict to subspaces $\mathcal{H}^+$ or $\mathcal{H}^-$ of $\psi \in \mathcal{H} \otimes \mathcal{H}$ that are either symmetric or antisymmetric under permutations implemented by the unitary operators,

$$U_\pi : \mathcal{H} \otimes \cdots \otimes \mathcal{H} \to \mathcal{H} \otimes \cdots \otimes \mathcal{H}$$

which permute factors according to the element $\pi \in \mathfrak{S}_N$. We require

$$U_\pi \psi = \begin{cases} \psi & \text{for } \psi \in \mathcal{H}^+ \\ \operatorname{sign}(\pi)\psi & \text{for } \psi \in \mathcal{H}^- \end{cases}$$

3 The Spectrum of Free N-Boson and N-Fermion Hamiltonians

Let $\Lambda \subseteq \mathbb{R}^d$ be a nice domain, e.g., $[0, L]^d \subseteq \mathbb{R}^d$. The one particle Hamiltonian is

$$H_1 = -\frac{\hbar^2}{2m} \triangle \tag{18}$$

with Dirichlet, Neumann, or periodic boundary conditions. The N particle Hamiltonian, H_N has to act on either the symmetric or antisymmetric subspace of $\bigotimes_{i=1}^N L^2(\Lambda)$. Suppose the eigenvalues (with repetitions in case of degeneracies) of H_1 are labeled by $k \in \Lambda^*$. Concretely, with periodic boundary conditions, we have $\Lambda^* = \frac{2\pi}{L}\mathbb{Z}^d/\text{mod } L$. Then

$$\varepsilon_k = \frac{\hbar^2}{2m}|k|^2 \tag{19}$$

are the eigenvalues of H_1, and

$$H_N = \sum_{i=1}^{N} \mathbb{1} \otimes \cdots \otimes \underbrace{H_1}_{i^{\text{th}}} \otimes \cdots \otimes \mathbb{1} \tag{20}$$

is the Hamiltonian acting on the appropriate subspace of $\bigotimes_{i=1}^N L^2(\Lambda)$.

Let $\varphi_k \in L^2(\Lambda)$ denote the eigenvector of H_1 belonging to the eigenvalue ε_k. The eigenvector of H_N on the full tensor product are simply the tensor products of the φ_k:

$$\psi_{k_1,\ldots,k_n} = \varphi_{k_1} \otimes \cdots \otimes \varphi_{k_n} \tag{21}$$

where $k_i \in \Lambda^*$, and $\varepsilon_{k_1,\ldots,k_n} = \sum_{i=1}^N \varepsilon_{k_i}$.

The Hamiltonians that we are interested in are of the form $P_N^\pm H_N P_N^\pm$, where $P_N^\pm$ is the orthogonal projection onto $\mathcal{H}_N^\pm$, the symmetric and antisymmetric subspaces of $\bigotimes_{i=1}^N L^2(\Lambda)$. Fortunately, since H_N is permutation invariant, we can use $P_N^\pm$ to turn the eigenvectors $\psi_{k_1,\ldots,k_n}$ into eigenvectors of $P_N^\pm$ (or zero) with the same eigenvalues, $\varepsilon_{k_1,\ldots,k_n}$. I.e., the non-zero vectors of the form

$$P_N^\pm \psi_{k_1,\ldots,k_n}$$

are eigenvectors of $P_N^\pm H_N P_N^\pm$.

Statistical Mechanics, Math 266: Week 9 Notes

March 2 and 4, 2010

1 Partition Functions for Ideal Fermi and Bose Gases

Let ε_k for $k \in \Lambda^*$ be the eigenvalues, with repetitions, of the one particle Hamiltonian in a bounded domain $\Lambda \subseteq \mathbb{R}^d$. The eigenvalues of the N particle Hamiltonians are

$$E_{k_1,\ldots,k_N} \qquad k_i \in \Lambda^* \text{ all distinct} \qquad (N \text{ Fermions})$$

$$E_{k_1,\ldots,k_N} \qquad (k_1,\ldots,k_N) \in (\Lambda^*)^N \qquad (N \text{ Bosons})$$

Defining $n_k = \#\{i | k_i = k \ , \ i = 1,\ldots,N\}$, we have that the eigenvalues will correspond to $(n_k)_{k\in\Lambda^*}$ such that $\sum_{k\in\Lambda^*} n_k = N$ and $n_k \geq 0$ and take integer values. Corresponding eigenvectors of the N particle Hamiltonian are anti-symmetric or symmetric with resepect to the labels $i = 1,\ldots,N$ of the tensor factors. Recall that the N particle Hamiltonian is

$$H_N^{\pm} P_N^{\pm} H_N P_N^{\pm} \tag{1}$$

where $H_N = \sum_{i=1}^N \mathbb{1} \otimes \cdots \otimes \mathbb{1} \otimes \underbrace{H_1}_{i^{\text{th}}} \otimes \mathbb{1} \otimes \cdots \otimes \mathbb{1}$, and $P_N^{\pm}$ is the orthogonal projection onto $\mathcal{H}^{\pm}$.

The canonical partition function is

$$Z^{\pm}(N,\beta) = \mathrm{Tr}_{\mathcal{H}_N^{\pm}}\, e^{-\beta H_N^{\pm}} \tag{2}$$

We also introduce the grand-canonical partition function, in which the particle number is not fixed:

$$Z^{\pm}(\Lambda,\beta,\mu) = \sum_{N\geq 0} \mathrm{Tr}_{\mathcal{H}_N^{\pm}}\, e^{-\beta(H^{\pm}-\mu N)} \tag{3}$$

For Fermions, the grand-canonical partition function takes the specific form

$$Z^{-}(\Lambda,\beta,\mu) = \sum_{n_k=0,1} e^{-\beta\sum_k(\varepsilon_k-\mu)n_k} \tag{4}$$

$$= \prod_{k\in\Lambda^*} (1 + e^{-\beta(\varepsilon_k-\mu)}) \tag{5}$$

and

$$\log(Z^-(\Lambda,\beta,\mu) = \mathrm{Tr}_{\mathcal{H}_1} \log(1+e^{-\beta(H_1-\mu\mathbb{1})}) \tag{6}$$

For Bosons,

$$Z^+(\Lambda,\beta,\mu) = \sum_{\{n_k\},\, n_k\geq 0} e^{-\beta\sum_{k\in\Lambda^*} n_k(\varepsilon_k-\mu)} \tag{7}$$

$$= \prod_{k\in\Lambda^*}\sum_{n\geq 0} e^{-\beta n(\varepsilon_k-\mu)} \tag{8}$$

$$= \prod_{k\in\Lambda^*}\frac{1}{(1-e^{-\beta(\varepsilon_k-\mu)})} \tag{9}$$

and

$$\log Z^+(\Lambda,\beta,\mu) = -\,\mathrm{Tr}_{\mathcal{H}_1}\log(1-e^{-\beta(H_1-\mu\mathbb{1})}) \tag{10}$$

So, to summarize:

$$\log Z^\pm(\Lambda,\beta,\mu) = \mp\,\mathrm{Tr}_{\mathcal{H}_1}\log(1\mp e^{-\beta(H_1-\mu\mathbb{1})}) \tag{11}$$

Similar to the *free energy* of an N particle system and its thermodynamic limit where we fix $\rho = \frac{N}{|\Lambda|}$, one defines the *pressure* of a grand-canonical system

$$p_\Lambda(\beta,\mu) = \frac{1}{\beta|\Lambda|}\log(Z(\Lambda,\beta,\mu)) \tag{12}$$

and

$$p(\beta,\mu) = \lim_{\Lambda\nearrow\mathbb{R}^d} p_\Lambda(\beta,\mu) \tag{13}$$

One then has $\frac{\partial p_\Lambda}{\partial\mu} = \frac{\omega(N)}{|\Lambda|}$, and in the limit $\Lambda\nearrow\mathbb{Z}^d$ or $\Lambda\nearrow\mathbb{R}^d$,

$$\frac{\partial p(\beta,\mu)}{\partial\mu} = \rho \tag{14}$$

relating the grand-canonical μ with the canonical parameter ρ.

2 The Pressure in the Thermodynamic Limit

In this section, we make the assumptions that $\Lambda = [0,L]^d \subseteq \mathbb{R}^d$ with periodic boundary conditions so that $\varepsilon_k = (\frac{2\pi}{L})^2 k^2$, and $k\in\mathbb{Z}^d$. For Fermions,

$$p(\beta,\mu) = \lim_{L\to\infty}\frac{1}{|\Lambda|\beta}\sum_{k\in\Lambda^*}\log\left[1+e^{-\beta\left((\frac{2\pi}{L})^2k^2-\mu\right)}\right] \tag{15}$$

For $\mu\geq 0$, this is a convergent Riemann sum. Hence,

$$p(\beta,\mu) = \frac{1}{(2\pi)^d\beta}\int_{-\infty}^{+\infty} d^du\,\log(1+e^{-\beta(u^2-\mu)}) \tag{16}$$

$$= \frac{\Omega_d}{(2\pi)^d\beta}\int_0^\infty dr\,r^{d-1}\log(1+e^{-\beta(r^2-\mu)}) \tag{17}$$

and so we can calculate the density function,

$$\rho(\mu) = \frac{\partial}{\partial \mu} p(\beta,\mu) = \frac{\Omega_d}{(2\pi)^d} \int_0^\infty dr\, r^{d-1} \frac{e^{-\beta(r^2-\mu)}}{1+e^{-\beta(r^2-\mu)}} \tag{18}$$

For Bosons, for convergence as $\Lambda \nearrow \mathbb{R}^d$, one has to assume $\mu < 0$ and then find in the same way that

$$p(\beta,\mu) = -\frac{\Omega_d}{(2\pi)^d \beta} \int_0^\infty dr\, r^{d-1} \log(1 - e^{-\beta(r^2-\mu)}) \tag{19}$$

and

$$\rho(\mu) = -\frac{\Omega_d}{(2\pi)^d} \int_0^\infty dr\, r^{d-1} \frac{1}{e^{\beta(r^2-\mu)} - 1} \tag{20}$$

If one is interested in the equilibrium state of the ideal Bose gas at a given β and particle density $\bar{\rho} \geq 0$, we look for a solution for μ of the equation

$$\rho(\mu) = \bar{\rho} \tag{21}$$

and then hope for an equivalence of ensembles result, as expected from the classical case, such that $\frac{N_\Lambda}{|\Lambda|} \to \rho$ and that the two expressions $\frac{e^{-\beta(H_\Lambda, N_\Lambda)}}{Z(\Lambda,\beta,N_\Lambda)}$ and $\bigoplus_{N\geq 0} \frac{e^{-\beta(H-\mu N)}}{Z(\Lambda,\beta,\mu)}$ converge to the same infinite volume state in the thermodynamic limit. Here, we set $\Lambda = \Lambda_L$ and choose N_L such that $\frac{N_L}{|\Lambda_L|} = \rho_L \to \bar{\rho}$. It turns out that as long as we find a solution $\mu < 0$ of (21), this approach indeed works. (Convergence details along these lines are discussed later).

In particular for $d = 1, 2$ this works. For $d = 1$,

$$\rho(\mu) = c \int_0^\infty dr \frac{1}{e^{\beta(r^2-\mu)} - 1} \tag{22}$$

For $\mu = 0$, by making the substitution $x = r^2$,

$$\rho(\mu) = \tilde{c} \int_0^\infty dx \frac{1}{\sqrt{x}} \frac{1}{e^{\beta x} - 1} = +\infty \tag{23}$$

Similarly, for $d = 2$, one finds that

$$\lim_{\mu\to 0} \rho(\mu) = c \int_0^\infty dx \frac{1}{e^{\beta x} - 1} = +\infty \tag{24}$$

(but the divergence is borderline). For $d = 3$, and in fact, for dimensions $d > 2$, we have

$$\lim_{\mu \nearrow 0} \rho(\mu) = \rho_c = \tilde{c} \int_0^\infty dx\, x^{\frac{d}{2}-1} \frac{1}{e^{\beta x} - 1} < +\infty \tag{25}$$

This however, does not mean that the ideal Bose gas in $d = 3$ does not have equilibrium points with densities $\bar{\rho} > \rho_c(\beta)$. Rather, the finiteness of the integral at $\mu = 0$ suggest that in *finite-volume* we consider $\mu_L > 0$. If $\mu_L \to \mu > 0$ as

$L \to \infty$ because of the singularities of $\frac{x^{\frac{d}{2}-1}}{e^{\beta(x-\mu)}}$ at $x = \mu$, which is not integrable. So, we will have to work with $\mu_L \to 0$. Recall that

$$\log Z_\Lambda(\beta, \mu) = \sum_{k \in \Lambda^*} \log(1 + e^{-\beta(\varepsilon_k - \mu)}) \tag{26}$$

Now, the details will depend on the chosen chosen boundary conditions. For periodic boundary conditions, $\Lambda_L = [0, L]^d$, $\varepsilon_k = \left(\frac{2\pi}{L}\right)^2 |k|^2$ with $k \in \mathbb{Z}^d$. Again, the pressure is

$$p_\Lambda(\beta, \mu) = \frac{1}{|\Lambda|\beta} \sum_{k \in \Lambda^*} \log\left(1 + e^{-\beta(\varepsilon_k - \mu)}\right) \tag{27}$$

and

$$\rho_\Lambda(\beta, \mu) = \frac{\partial}{\partial \mu} p_\Lambda(\beta, \mu) \tag{28}$$

$$= \frac{1}{|\Lambda|} \sum_{k \in \Lambda^*} \frac{1}{e^{-\beta(\varepsilon_k - \mu)} - 1} \tag{29}$$

$$= \frac{e^{\beta\mu}}{1 - e^{\beta\mu}} + \frac{1}{|\Lambda|} \sum_{k \in \Lambda^*} \frac{1}{e^{\beta\varepsilon_k} - e^{\beta\mu}} \tag{30}$$

This last sum converges if $\mu_{|\Lambda|} \to 0$. If $\bar{\rho} > \rho_c$, pick $\mu_{|\Lambda|}$ such that

$$|\Lambda|(\bar{\rho} - \rho_c) = \frac{1}{e^{-\beta\mu_{|\Lambda|}} - 1}$$

or

$$1 + \frac{1}{|\Lambda|(\bar{\rho} - \rho_c)} = e^{-\beta\mu_{|\Lambda|}}$$

and so solving for $\mu_{|\Lambda|}$ gives

$$\mu_{|\Lambda|} = -\frac{1}{\beta} \log\left(1 + \frac{1}{|\Lambda|(\bar{\rho} - \rho_c)}\right) \tag{31}$$

$$\sim -\frac{1}{\beta(\bar{\rho} - \rho_c)|\Lambda|} \nearrow 0 \tag{32}$$

What about the pressure?

$$p_\Lambda(\beta, \mu_{|\Lambda|} = \frac{1}{|\Lambda|\beta} \sum_{k \in \Lambda^*} \log\left(1 + e^{-\beta(\varepsilon_k - \mu_{|\Lambda|})}\right) \tag{33}$$

and it is not hard to show that $p_\Lambda(\beta, \mu_{|\Lambda|}) \to p(\beta, 0)$. It now follows that the pressure and free energy will be constant for $\rho > \rho_c$, and will be equal to the value for $\rho = \rho_c$.

$$f(\beta, \rho) = \sup_{\mu \leq 0}(\mu\rho - p(\beta, \mu)) \tag{34}$$

3 Properties of the Grand Canonical Equilibrium State

Let $N_{0,\Lambda}$ denote the number of particles in the $k = 0$ state. Then

$$N_{0,\Lambda} = \frac{\sum_{n_0=0}^{\infty} n_0 e^{\beta\mu n_0} \sum_{\substack{n_k \\ 0\neq k\in\Lambda^*}} e^{-\beta\sum_{k\in\Lambda^*} n_k(\varepsilon_k-\mu)}}{Z^+(\Lambda,\beta,\mu)} \tag{35}$$

$$= \frac{\frac{e^{\beta\mu}}{(1-e^{\beta\mu})^2}\prod_{0\neq k\in\Lambda^*}(1-e^{-\beta(\varepsilon_k-\mu)})^{-1}}{\prod_{k\in\Lambda^*}(1-e^{-\beta(\varepsilon-\mu)})^{-1}} \tag{36}$$

$$= \frac{e^{\beta\mu}}{1-e^{\beta\mu}} \tag{37}$$

which is the same equation as derived for $\mu_{|\Lambda|}$. So as expected (by design), we have

$$N_{0,\Lambda} = |\Lambda|(\bar{\rho}-\rho_c) \tag{38}$$

and the condensate density is

$$\frac{N_{0,\Lambda}}{|\Lambda|} \to \bar{\rho}-\rho_c \tag{39}$$

For $k \neq 0$, $N_k \sim L^2$, and so

$$\frac{N_{k,\Lambda}}{|\Lambda|} \to 0 \tag{40}$$

and for $\varepsilon > 0$,

$$N_\varepsilon \sim \frac{e^{-\beta\varepsilon}}{(1-e^{-\beta\varepsilon}} < \infty \tag{41}$$

This condensation of a finite fraction of the particles in the minimum energy state is called *Bose-Einstein condensation*. Next, we shall see how this phenomena is accompanied by the spontaneous breaking of a continuous symmetry.

Statistical Mechanics, Math 266: Week 9 Notes, Part 2

March 4, 2010

1 Second Quantization Formalism and Continuous Symmetry Breaking in the Ideal Bose Gas

Our definition of $P_N^{\pm}$ can be extended to any n-fold tensor product of a Hilbert space $\mathcal{H}$

$$\mathcal{H}_N = \bigotimes_{i=1}^{N} \mathcal{H}$$
$$\mathcal{H}_N^{\pm} = P_N^{\pm}\mathcal{H}_N$$
$$\mathcal{F}^{\pm} = \bigoplus_{N\geq 0} \mathcal{H}_N^{\pm}$$

where F^+ corresponds to Bosonic Fock space and F^- corresponds to Fermionic Fock space with one particle space $\mathcal{H}$. For all A linear operators defined on $\mathcal{H}$,

$$A^{(N)} = A \otimes \mathbb{1} \otimes \cdots \otimes \mathbb{1} + \cdots \mathbb{1} \otimes \cdots \otimes \mathbb{1} \otimes A$$

and

$$P_N^{\pm}\varphi_1 \otimes \cdots \otimes \varphi_N = \frac{1}{N!}\sum_{\pi\in\mathfrak{S}_N} (\pm 1)^{\operatorname{sign}\pi} \varphi_{\pi(1)} \otimes \cdots \otimes \varphi_{\pi(N)}$$

leave $\mathcal{H}_N^{\pm}$ invariant and thus are well-defined when restricted to $\mathcal{H}_N^{\pm}$. The second quantization of A is then

$$d\Gamma(A) = \bigoplus_{N\geq 0} A^{(N)} \text{ on } \mathcal{F}^{\pm}$$

The creation and annihilation operators $a^+(f) : \mathcal{H}_N^{\pm} \to \mathcal{H}_{N+1}^{\pm}$ and $a(f) : \mathcal{H}_{N+1}^{\pm} \to \mathcal{H}_N^{\pm}$ are defined on $\mathcal{F}^{\pm}$ for all $N \geq 0$, where $\mathcal{H}_N^{\pm}$ and $\mathcal{H}_{N+1}^{\pm}$ are considered as subspaces of $\mathcal{F}^{\pm}$. They are defined as follows:

$$a^+(f)P_N^{\pm}(\varphi_1 \otimes \cdots \otimes \varphi_N) = \sqrt{N+!}P_{N+1}^{\pm}(f \otimes \varphi_1 \otimes \cdots \otimes \varphi_N) \tag{1}$$

$$a(f)P_N^{\pm}(\varphi_1 \otimes \cdots \otimes \varphi_N) = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} (\pm 1)^{k-1} \langle f, \varphi_k \rangle P_{N-1}^{\pm}(\varphi_1 \otimes \cdots \otimes \hat{\varphi}_k \otimes \cdots \varphi_N) \tag{2}$$

where $\hat{\varphi}_k$ denotes that the entry φ_k is missing from the tensor product.

Important properties of the creation and annihilation operators (as proved in the homework and in Bratelli and Robinson, Volume 2) are:

1. $a^+(f) = a(f)^*$, when defined on their natural domains in $\mathcal{F}^{\pm}$
2. $a^+(f)$ depends linearly on f, and $a(f)$ depends anti-linearly
3. for Bosons,
$$[a(f), a^+(g)] \subseteq \langle f, g \rangle \mathbb{1} \tag{3}$$
for Fermions,
$$\{a(f), a^+(g)\} = \langle f, g \rangle \mathbb{1} \tag{4}$$
4. Define $\mathbf{N} = d\Gamma(\mathbb{1})$. Then $\mathbf{N}|_{\mathcal{H}_N^{\pm}} = N\mathbb{1}$.
5. Let (φ_k) be an orthonormal basis of $\mathcal{H}$. Then
$$\mathbf{N} = \sum_k a^+(\varphi_k) a(\varphi_k) \tag{5}$$
6. if $A = \sum_{k,l} |\varphi_k\rangle A_{k,l} \langle \varphi_l|$, then
$$d\Gamma(A) = \sum_{k,l} A_{k,l} a^+(\varphi_k) a(\varphi_l) \tag{6}$$
7. For Fermions,
$$\|a^+(f)\| = \|a(f)\| = \|f\| \text{ for all } f \in \mathcal{H} \tag{7}$$

With H_N as before,

$$\bigoplus_N H_N = d\Gamma(H_1) \tag{8}$$

and so the grand canonical density matrix on $\mathcal{F}$ for $\mathcal{H} = L^2(\Lambda)$ is defined by

$$\rho_{\beta,\mu} = \frac{1}{Z} e^{-\beta(\mathbf{H} - \mu \mathbf{N})} \tag{9}$$

where $\mathbf{H} = d\Gamma(H)$. Let $\{\varphi_k\}$ be an eigenbasis of H_1 as before with eigenvalues ε_k. The number operator $\mathbf{N}$ is a densely defined self-adjoint operator with domain $D(N) = \{\psi = \bigoplus_N \psi_N \in \mathcal{F}^{\pm} | \sum_N N^2 \|\psi_N\|^2 < +\infty\}$ and so for all

$\alpha \in \mathbb{R}$, $U(\alpha) = e^{i\alpha\mathbf{N}}$ is a unitary operator on Fock space. $e^{i\alpha} \mapsto U(\alpha)$ is a representation of $U(1)$. The second quantized Hamiltonian $\mathbf{H}$ has the gauge symmetry

$$[\mathbf{H}, U(\alpha)] = 0 \tag{10}$$

which expresses that $\mathbf{H}$ conserves particle number. Its infinitesimal form is $[\mathbf{H}, \mathbf{N}] = 0$. This is the continuous symmetry that is broken in the Bose-Einstein condensed phase of the ideal Bose gas. Before we explain this in detail , we first introduce the reduced density matrices. The one-particle reduced density matrix (which is *not* the density matrix of the one-particle system) is defined as the operator $\rho^{(1)}$ with matrix elements

$$\langle g, \rho^{(1)} f\rangle_{\beta,\mu} = \omega_{\beta,\mu}(a^+(f)a(g)) \tag{11}$$

and the finite volume density matrix is given by

$$\rho_{\beta,\mu} = \frac{e^{-\beta\sum_{k\in\Lambda^*}(\varepsilon_k-\mu)a^+(\varphi_k)a(\varphi_k)}}{Z(\beta,\mu)}$$

$$\begin{aligned}
\langle\varphi_l, \rho^{(1)}\varphi_k\rangle &= \omega_{\beta,\mu}(a^+(\varphi_k)a(\varphi_l)) \\
&= \frac{1}{Z}\sum_{n_k}\sum_{n_l} e^{-\beta\sum_k n_k(\varepsilon_k-\mu)}\langle n_k a^+(\varphi_k)a(\varphi_l)n_l\rangle\delta_{k,l}n_k \\
&= \omega_{\beta,\mu}(N_k) = \frac{e^{-\beta(\varepsilon_k-\mu)}}{1-e^{-\beta(\varepsilon_k-\mu)}} \text{ (The Bose distribution function)} \\
&= \frac{1}{e^{\beta(\varepsilon_k-\mu)}-1}
\end{aligned}$$

and thus,

$$\rho^{(1)} = \sum_k \frac{1}{e^{-\beta(\varepsilon_k-\mu)}-1}|\varphi_k\rangle\langle\varphi_k| \tag{12}$$

This operator will have a kernel $\rho^{(1)}(x,y)$,

$$\rho^{(1)}(x,y) = \sum_{k\in\Lambda^*}\frac{1}{e^{\beta(\varepsilon_k-\mu)}-1}\frac{1}{|\Lambda|}e^{ik(x-y)}$$

where $k \in (\frac{2\pi}{L})\mathbb{Z}^d$ and we have made use of the fact that $\varphi_k = \frac{1}{\sqrt{|\Lambda|}}e^{ikx}$ and as $\Lambda \nearrow \mathbb{R}^d$,

$$\rho^{(1)}(x,y) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} dk\frac{e^{ik(x-y)}}{e^{\beta(k^2-\mu)}-1}$$

This derivation is again valid for fixed $\mu < 0$ or $\mu_L \to \mu_\infty < 0$. If $\mu_L \to 0$, we need to separate out the $\mu = 0$ torm as we did before for the calculation of N_0. So, for $\bar{\rho} > \rho_c$ the result becomes

$$\rho^{(1)}(x,y) = (\bar{\rho}-\rho_c)|1\rangle\langle 1| + \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} dk\frac{e^{ik(x-y)}}{e^{\beta k^2}-1}$$

Note that $\rho^{(1)}(x,y) = \bar{\rho}(x-y)$. By Riemann-Lebesgue for $\bar{\rho} \le \rho_c$,

$$\lim_{r\to\infty} \sigma(r) = 0$$

but for $\bar{\rho} > \rho_c$,

$$\lim_{r\to\infty} \sigma(r) = \bar{\rho} - \rho_c > 0$$

This is called off-diagonal long range order (OLRO, Yang). (Think of $\rho(x,y)$ as a matrix indexed by x). For comparison, consider LRO in the Ising model where

$$\omega = \frac{1}{2}(\omega_+ + \omega_-)$$

is spin-flip symmetric but does not cluster, meaning

$$\omega(\sigma_x) = 0 \text{ for all } x \in \mathbb{Z}^d \ (d \ge 2)$$

$$\omega_\pm(\sigma_x) = \pm m(\beta) \text{ where } 0 < m(\beta) \text{ for } \beta > \beta_c$$

One can also show that in these states

$$\omega_\pm(\sigma_x \sigma_y) = \sigma_\pm(x-y)$$

$$\lim_{r\to\infty} \omega_\pm(\sigma_0\sigma_r) = m^2 = \omega_\pm(\sigma_0)\omega_\pm(\sigma_r)$$

so that these states exhibit clustering, but

$$\lim_{r\to\infty} \omega(\sigma_0\sigma_r) = m^2$$

as well, but

$$\lim_{r\to\infty} \omega(\sigma_0\sigma_r) \ne \omega(\sigma_0)\omega(\sigma_r) = 0$$

so that ω does not cluster and this has long range order. To complete the analogy we need to construct the equilibrium states with *broken* gauge symmetry. Instead of boundary conditions, we will achieve this by adding an "external field" to the Hamiltonian and then taking the limit of zero-field strength.

Let $a_0^{(*)}$ denote $a^{(*)}(\varphi_0)$ where $\varphi_0(x) = \frac{1}{|\Lambda|^{\frac{1}{2}}}\chi_\Lambda(x)$.

$$H_\Lambda^{(\lambda)} = H_\Lambda + \lambda|\Lambda|^{\frac{1}{2}} a_0^* + \bar{\lambda}|\Lambda|^{\frac{1}{2}} a_0 \tag{13}$$

Now, we calculate the λ-dependent quantities

$$p_\Lambda(\beta,\mu,\lambda) = \frac{1}{\beta|\Lambda|} \log \mathrm{Tr}[e^{-\beta(H_\lambda^{(\Lambda)} - \mu \mathbf{N})}]$$

Define $B(f) = a^*(f) + a(f)$ (the operator theoretic closure of $B(f)$ is a self-adjoint operator with denso domain $D(B(f))$) and then $W(f) = e^{iB(f)}$ is unitary. The $W(f)$ are called Weyl operators. They satisfy the following important identities which all follow from the CCR $[a(f), a^+(g)] = \langle f, g\rangle \mathbb{1}$:

$$[a(f), W(g)] = i\langle f, g\rangle W(g) \tag{14}$$

$$[a^+(f), W(g)] = -i\langle \bar{f, g}\rangle W(g) \tag{15}$$

$$[B(f), W(g)] = -2\Im\langle f, g\rangle W(g) \tag{16}$$

$$W(-g)B(f)W(g) = B(f) - 2\Im\langle f, g\rangle \tag{17}$$

$$W(-g)W(g) = \mathbb{1} \tag{18}$$

$$W(-g)W(f)W(g) = e^{-2i\Im\langle f,g\rangle} W(f) \tag{19}$$

These last implications follow by noting

$$W(-g)\,(iB(f))^n\,W(g) = (W(-g)(iB(f))W(g))^n$$

for all $n \geq 0$ and finally, the famous Weyl relations

$$W(f)W(g) = e^{-i\Im\langle f,g\rangle} W(f+g) \tag{20}$$

To prove the latter, consider the following derivative with respect to a real variable t:

$$\begin{aligned}
\frac{d}{dt}[W(tf)W(tg)W(-t(f+g))] &= W(tf)iB(f)W(tg)W(-t(f+g)) \\
&+ W(tf)W(tg)(iB(g))W(-t(f+g)) \\
&+ W(tf)W(tg)(-i(B(f)+B(g)))W(-t(f+g)) \\
&= W(tf)[iB(f), W(tg)]W(-t(f+g)) \\
&= -2it\Im\langle f, g\rangle W(tf)W(tg)W(-t(f+g))
\end{aligned}$$

Calling $V(t) = W(tf)W(tg)W(-t(f+g))$, and noting that $V(0) = \mathbb{1}$,

$$W(t)W(g)W(-(f+g)) = V(1) = \mathbb{1} - 2i\int_0^1 dt t\Im\langle f, g\rangle V(t)$$

Now by iterating,

$$\begin{aligned}
V(1) = \mathbb{1} &= i\Im\langle f, g\rangle + (-i\Im\langle f, g\rangle)^2 \int_0^1 dt\frac{t^2}{2} + \ldots \\
&= e^{-i\Im\langle f,g\rangle}
\end{aligned}$$

$$\begin{aligned}
W(-g)a^*(f)W(g) &= W(-g)[a^*(f), W(g)] + a^*(f) \\
&= a^*(f) - i\langle g, f\rangle
\end{aligned}$$

Put $g = \frac{-i\lambda\chi_\Lambda}{\varepsilon_0 - \mu}$. In the case of periodic boundary conditions, $g = i\frac{\lambda}{\mu}\chi_\Lambda$. Then

$$W(-g)a^*(\varphi_k)W(g) = a^*(\varphi_k) + \delta_{k,0}\frac{\bar{\lambda}|\Lambda|^{\frac{1}{2}}}{\varepsilon_0 - \mu}$$

and thus,

$$W(-g)(\mathbf{H}-\mu\mathbf{N})W(g) = \sum_{k\in\Lambda^*}(\varepsilon_k-\mu)a^*(\varphi_k)a(\varphi_k)+\bar{\lambda}|\Lambda|^{\frac{1}{2}}a(\varphi_0)+\lambda|\Lambda|^{\frac{1}{2}}a^*(\varphi_0)+\frac{|\lambda|^2|\Lambda|}{(\varepsilon_0-\mu)^2}$$

In other words, $H_\Lambda^{(\lambda)}$ is unitarily equivalent to $H_\Lambda^{(0)} - |\lambda|^2|\Lambda|\frac{1}{\mu^2}$, where we have assumed that $\varepsilon_0 = 0$. The constant does not affect $\rho_{\beta,\mu}$. This implies that

$$\rho_{\beta,\mu,\lambda} = W(-g)\rho_{\beta,\mu}W(g)$$

In particular, (all in finite volume Λ),

$$\begin{aligned}
\omega_{\beta,\mu,\lambda}(a^*(\varphi_k)) &= \omega_{\beta,\mu,0}(W(g)a^*(\varphi_k)W(-g)) \\
&= \omega_{\beta,\mu,0}(a^*(\varphi_k)) - \delta_{k,0}\frac{\bar{\lambda}|\Lambda|^{\frac{1}{2}}}{\mu} \\
&= -\delta_{k,0}\frac{\bar{\lambda}|\Lambda|^{\frac{1}{2}}}{\mu}
\end{aligned}$$

Now, we need to redo the calculation of $\mu_\Lambda(\lambda)$ for $\lambda \neq 0$.

$$\begin{aligned}
p_\Lambda(\mu,\beta,\lambda) &= \frac{1}{\beta|\Lambda|}\log \operatorname{Tr} e^{-\beta(H_\Lambda^{(\lambda)}-\mu N)} \\
&= p_\Lambda(\mu,\beta) + \frac{|\lambda|^2}{\mu}
\end{aligned}$$

Then

$$\frac{\langle a_0\rangle}{|\Lambda|^{\frac{1}{2}}} = \frac{\partial p}{\partial \lambda} = \frac{\lambda}{-\mu}$$

and

$$\frac{\langle N\rangle}{|\Lambda|} = \frac{\partial p}{\partial \mu}(\mu,\lambda) = \rho_\Lambda(\mu) + \frac{|\lambda|^2}{\mu^2}$$

Solving

$$\bar{\rho} = \rho_\Lambda(\mu) + \frac{|\lambda|^2}{\mu^2}$$

if $\bar{\rho} > \rho_c$, then $\frac{|\lambda|}{\mu(\lambda)} \to \sqrt{\bar{\rho}-\rho_c}$ and

$$\left|\frac{|\langle a_0\rangle_{\lambda,\mu}}{|\Lambda|^{\frac{1}{2}}}\right| \to \left|\frac{\lambda}{\mu}\right| \to (\bar{\rho}-\rho_c)^{\frac{1}{2}}$$

independent of Λ, which survives the limit $\lambda \searrow 0$. What we have done is constructed a symmetry-broken phase in which

$$|\omega_{\beta,\mu,0}\left(\frac{a_0}{|\Lambda|^{\frac{1}{2}}}\right) = \sqrt{\bar{\rho}-\rho_c}$$

by $\lambda \to e^{i\alpha}\lambda$ one gets a family of states parameterized by α such that

$$\omega_{\beta,\mu,0}\left(\frac{a_0}{|\Lambda|^{\frac{1}{2}}}\right) = e^{i\alpha}\sqrt{\bar{\rho} - \rho_c}$$

for all $\alpha \in [0, 2\pi)$. The symmetry that relates these states is $e^{i\alpha\mathbf{N}}$. Note that $[a_0^+ a_0, a_0] = -a_0$, and therefore

$$\begin{aligned} e^{-i\alpha\mathbf{N}} a_0 e^{i\alpha\mathbf{N}} &= e^{-i\alpha[\mathbf{N},\cdot]} a_0 \\ &= e^{i\alpha[a_0^+ a_0,\cdot]} a_0 \\ &= \sum_{n\geq 0} \frac{-i\alpha)^n}{n!}(-1)^n a_0 \\ &= e^{i\alpha} a_0 \text{ (gauge transformation)} \end{aligned}$$

Statistical Mechanics, Math 266: Week 10 Notes

March 9 and 11, 2010

1 A Model for Superfluidity due to Bogoliubov

The method of Bogoliubov Transformations, or Canonical Transformations, is related more generally to the Hartree-Fock and generalized Hartree-Fock model for high density Bose gases in the regime where $\rho \gg \rho_{\text{initial}}$, $\mu = 0$ where in fact most particles are assumed to be in the 0-state:

$$\rho_c = \frac{N - N_0}{V} \ ; \ N - N_0 \ll N_0$$

We also assume that the interaction potential has a positive Fourier Transform.

$$H = \sum_k \varepsilon_k a_k^+ a_k + \frac{1}{2V} \sum_{\substack{k,p \\ Q}} \hat{V}(Q) a_{k+Q}^+ a_{k-Q}^+ a_p a_k \tag{1}$$

with $\hat{V}(Q) \geq 0$ and $N_0 = \langle a_0^+ a_0 \rangle$. Bogoliubov's insight was that in this regime, behavior will be well described by a simplified Hamiltonian where we "only keep interation terms with two or more factors of a_0 (of order N_0 or higher)".

$$H_1 = \sum_{k \neq 0} (\varepsilon_k + \frac{\hat{V}(k)}{|\Lambda|} N) a_k^+ a_k + \frac{\hat{V}(0)}{2|\Lambda|} N^2 + \frac{1}{2} N \sum_{k \neq 0} \hat{V}(k) (a_k a_{-k} + a_k^+ a_{-k}^+) \tag{2}$$

There are a wide variety of models (Hamiltonians) of this form. We will illustrate in this simple case and at $T = 0$ only that Bogoliubov Transformations can be used to find the ground state (equilibrium state) of such models. (Also for Fermions (BCS Theory)).

2 Bogoliubov (Canonical) Transformations

Canonical transformations that leave the commutation relations invariant (CAR or CCR), i.e. they are automorphisms of the algebras generated by the CAR and CCR. In analogy with classical mechanics, the name *canonical transformations* is reserved for the transformation of the underlying phase space , a.k.a., the one-particle space. In the case of the CCR, this space has a symplectic structure and the canonical transformation can be viewed as the second quantized form of symplectic transforms on that space.

Let's proceed somewhat informally. A BCT (Bogoliubov Canonical Transformation) can then be regarded as an automorphism of the CCR (or CAR) of the following form:

$$b(f) = \tau(a(f)) = a(Uf) + a^+(Vf) \text{ , for } f \in \mathcal{H}$$

where U is a linear operator on $\mathcal{H}$ and V is an antilinear operator on $\mathcal{H}$. (Antilinear means $V(f+\lambda g) = V(f) + \bar{\lambda}V(g)$ for $f \in \mathcal{H}$ and $\lambda \in \mathbb{C}$ and its adjoint V^* is an *antilinear* operator determined by $\langle f|Vg\rangle = \overline{\langle V^* f|g\rangle}$. Since τ is supposed to be an automorphism, the $b(f)$ satisfy the same algebraic relations as the $a(f)$ (the CCR in our case). For this to hold, U and V have to satisfy the following relations:

$$\begin{aligned} U^*U = V^*V = \mathbb{1} = UU^* - VV^* \\ U^*V - V^*U = 0 = UV^* - VU^* \end{aligned} \tag{3}$$

automorphisms ar invertible and using the relations in (3) one can show that τ^{-1} is given by

$$\tau^{-1}(a(f)) = a(U^* f) - a^+(V^* f)$$

For example, we can check that with $b(f) = \tau(a(f))$ and $b^+(f) = \tau(a^+(f)) = \tau(a(f))^*$ that the $b(f)$ indeed satisfy the CCR:

$$[b(f), b^+(g)] = \langle f, g\rangle \mathbb{1} \text{ for all } f, g \in \mathcal{H}$$
$$[b(f), b(g)] = [b^+(f), b^+(g)] = 0$$

Remark 2.1. *Antilinear operators V are typically given in terms of a linear operator V_1 and the complex conjugation operator:*

$$Vf = V_1\bar{f}$$

Since $b(f), b^+(f)$ satisfy the CCR we can think of them as creation and annihilation operators of a new kind of particle, the so-called concept of a quasi-particle. This is one way in which the quantum theory of particles is really different from the classical theory. Another novel notion in quantum physics is that the vacuum state is defined with respect to a particular notion of particles (the state with no such particles). The vacuum is however, not just nothing. In particular, the vacuum state for one kind of particles is not the same as the vacuum as the vacuum for quasi-particles defined using a non-trivial BCT.

For simplicity we will here only discuss the application of BCTs to finding ground states ($T = 0$) of quadratic Hamiltonians. More generally, BCTs are used to carry out the Hartree-Fock approximation at both $T = 0$ and $T > 0$. We concentrate on $T = 0$ here. The main idea is the following: Find τ (i.e. U, V) such that in terms of the $b(f) = \tau(a(f))$ H can be written in the form

$$H = \sum_\alpha \omega_\alpha b^+(f_\alpha)b(f_\alpha) + E_0$$

with $\omega_\alpha \geq 0$ and $E_0 \in \mathbb{R}$. Then E_0 is the ground state energy of H and the ground state is the vacuum (quasi-vacuum) for the (quasi-)particles $b(f)$. If this can be done, one is actually doing an exact diagonalization of H. Clearly, only H's that are at most quadratic in the $a(f)$'s will allow such a treatment (τ does not change the order of the highest order terms).

We will apply a BCT to Bogoliubov's Hamiltonian:

$$H = \sum_{k \neq 0} (\varepsilon_k + \frac{\hat{V}(k)}{|\Lambda|} N) a_k^+ a_k + \frac{\hat{V}(0) N^2}{2|\Lambda|} + \frac{1}{2} N \sum_{k \neq 0} \hat{V}(k)(a_k a_{-k} + a_k^+ a_{-k}^+)$$

with $\varepsilon_k = k^2$, $\hat{V}(k) \geq 0$, N is the number of particles, and $\rho = \frac{N}{|\Lambda|}$. As before, let $\Lambda = [0, L]^d \subseteq \mathbb{R}^d$, $\varepsilon_k = k^2$, $k \in \Lambda^*$, $\varphi_k = \frac{e^{ik \cdot x}}{|\Lambda|^{\frac{1}{2}}}$ then $\overline{\varphi_k} = \varphi_{-k}$. We will consider maps U and V that are "diagonal" in the basis φ_k:

$$U \varphi_k = u_k \varphi_k$$
$$V \varphi_k = -v_k \varphi_k$$

and assume $u_{-k} = u_k$ and $v_{-k} = v_k$ where $u_k, v_k \in \mathbb{R}$.

$$b_k = u_k a_k - v_k a_{-k}^+ \tag{4}$$

and one can check that

$$U^* U - V^* V = \mathbb{1} = U U^* - V V^*$$

holds if and only if

$$u_k^2 - v_k^2 = 1 \tag{5}$$

(4) of course implies

$$b_k^+ = u_k a_k^+ + v_k a_{-k}$$

and the inverse relations are easy to obtain using (5).

$$a_k = u_k b_k + v_k b_{-k}^+$$
$$a_k^+ = u_k b_k^+ + v_k b_{-k}$$

Now plug this into the expression of Bogoliubov's model Hamiltonian H_1 and simplify. One obtains:

$$H_1 = \Delta + G_0 + G_1 \tag{6}$$

where

$$\Delta = \frac{1}{2} \rho N \hat{V}(0) + \sum_{k \neq 0} \{(\varepsilon_k + \rho \hat{V}(k)) v_k^2 + \rho \hat{V}(k) u_k v_k\} \tag{7}$$

$$G_0 = \sum_{k \neq 0} \{\underbrace{(\varepsilon_k + \rho \hat{V}(k))(u_k^2 + v_k^2) + 2\rho \hat{V}(k) u_k v_k}_{\equiv \omega_k \geq 0}\} b_k^+ b_k \tag{8}$$

$$G_1 = \sum_{k \neq 0} \{(\varepsilon_k + \rho \hat{V}(k)) u_k v_k + \frac{1}{2} \rho \hat{V}(k)(u_k^2 + v_k^2)\}(b_k^+ b_{-k}^+ + b_k b_{-k}) \tag{9}$$

We will achieve our goal if we can choose u_k, v_k such that G_1 vanishes, i.e.,

$$0 = 2(\varepsilon_k + \rho \hat{V}(k)) u_k v_k + \rho \hat{V}(k)(u_k^2 + v_k^2) \text{ for all } k \in \Lambda^* \tag{10}$$

The condition $u_k^2 - v_k^2 = 1$ can be satisfied by (without loss of generality) taking u_k and v_k of the form:

$$\begin{aligned} u_k &= \varepsilon_k \cosh \alpha_k \\ v_k &= \eta_k \sinh \alpha_k \end{aligned}$$

where $\varepsilon_k, \eta_k = \pm 1$ and $\alpha_k \geq 0$. In this parameterization,

$$\begin{aligned} u_k v_k &= \varepsilon_k \eta_k \cosh \alpha_k \sinh \alpha_k = \frac{1}{2} \varepsilon_k \eta_k \sinh 2\alpha_k \\ u_k^2 + v_k^2 &= \cosh^2 \alpha_k + \sinh^2 \alpha_k = \cosh 2\alpha_k \end{aligned}$$

The condition (10) then becomes

$$0 = (\varepsilon_k + \rho \hat{V}(k)) \varepsilon_k \eta_k \sinh 2\alpha_k + \rho \hat{V}(k) \cosh 2\alpha_k$$

i.e.,

$$\tanh 2\alpha_k = \frac{\rho \hat{V}(k)}{\varepsilon_k + \rho \hat{V}(k)}$$

where we recall our assumption that $\hat{V}(k) \geq 0$, and pick $\eta_k = -1$ and $\varepsilon_k = +1$ for all k. Now, using the facts that

$$\sinh 2\alpha = \tanh 2\alpha \frac{1}{\sqrt{1 - (\tanh 2\alpha)^2}}$$

and

$$\cosh^2 2\alpha - \sinh^2 2\alpha = 1$$

we obtain

$$\sinh 2\alpha_k = \frac{\rho \hat{V}(k)}{\sqrt{(\varepsilon_k + \rho \hat{V}(k))^2 - \rho^2 \hat{V}(k)^2}} \tag{11}$$

$$\cosh 2\alpha_k = \frac{\varepsilon_k + \rho \hat{V}(k)}{\sqrt{(\varepsilon_k + \rho \hat{V}(k))^2 - \rho^2 \hat{V}(k)^2}} \tag{12}$$

and

$$\omega_k = \sqrt{(\varepsilon_k + \rho \hat{V}(k))^2 - \rho^2 \hat{V}(k)^2}$$

In other words,

$$H_1 = \sum_{k \neq 0} \omega_k b_k^+ b_k + \Delta \tag{13}$$

and the ground state is the vacuum for the b_k.

It is interesting to analyze the behavior of ω_k:

$$\varepsilon_k = k^2$$

so for large k and assuming $\hat{V}(k) \to 0$ for large k, (e.g., $V \in L^1$), then we also have $\omega \sim k^2$.

If we assume $\hat{V}(0) > 0$ and $\hat{V}(k)$ is continuous near $k = 0$ then for small k,

$$\omega_k \sim \sqrt{2\rho\hat{V}(0)}|k|$$

In particular, with the details depending on the specifics of $\hat{V}(k)$ we may now have a non-convex, non-monotone dispersion relation ω_k:

We can see this behavior in an explicit example. Consider $V(x) = e^{-\gamma|x|}$ where $\gamma > 0$.

$$\hat{V}(k) = \frac{\hat{V}(0)\gamma}{|k|^2 + \gamma}$$

$$\omega_k = \sqrt{\frac{(|k|^2(|k|^2+\gamma) + \rho\hat{V}(0)\gamma)^2 - \rho^2\hat{V}(0)\gamma^2}{(|k|^2+\gamma)^2}}$$

Then for γ not too large, ω_k looks like the graph above.

An interesting consequence of this type of dispersion relation, the point of the Bogoliubov model, is that it predicts superfluidity. The basic idea is as follows:

A normal fluid experiences resistance and friction as it flows because perturbations may extract kinetic energy out of the flow. In quantum mechanics, we expect this to happen if the "perturbation" sees an excited state it couples to that is a state of "extra energy" with respect to the ground state. This is of course generically the case, but with ω_k as in the figure, the mere fact that an observer (perturbating agent) sees the fluid in uniform motion, is not enough to have negative energy perturbations of the flowing fluid, or any if the velocity is not too high. That is, for all k, the flowing fluid still has $\tilde{\omega}_k \geq 0$ as seen by a moving observer. Thus, transfering momentum would mean creating a $\tilde{b}_k$ with $\tilde{\omega}_k > 0$, i.e., adding rather than extracting energy of the fluid. Thus, there is no dissipation.

$$\tilde{\omega}_k = \omega_k + k \cdot v$$

where v is the velocity of the uniform motion. Transfer of momentum k to fluid implies change in momentum $-k$ to the environment and the change in energy can then be estimated to be

$$\Delta E = \frac{1}{2M}(Mv_e - k)^2 - \frac{1}{2}Mv_e^2$$

where v_e is the velocity of the environment, $M \gg 1$ is the mass of the environment.

$$\Delta E = \frac{k^2}{2M} - k \cdot v_e \sim -k \cdot v_e$$

and $v_e = -v$ if v is the speed of uniform motion of the fluid described by our model. Therefore,

$$\tilde{\omega}_k = \omega_k + k \cdot v$$

and if $v < v_c$, $\tilde{\omega}_k \geq 0$ and there is no dissipation, hence there is superfluidity. k_c where $\omega_{k_c} = k_c \cdot v_c$ predicts the momentum of the first instability in the superfluid when it reaches the critical velocity v_c.