Tomás Pérez
Condensed Matter Theory - Lecture Notes
October 24, 2022

Theory & Notes

1. QUANTUM PHASE TRANSITIONS

Consider a Hamiltonian $\mathbf{H}(g)$, whose degrees of freedom reside on the sites of a lattice $\Lambda \subseteq \mathbb{Z}^d$, and which varies a function of a dimensionless coupling g. Then,

- for the case of a finite lattice, the ground state energy will be, in general, a smooth, analytic function of g. The main exception comes from the case when g couples only to a conserved quantity, ie.

$$\mathbf{H}(g) = \mathbf{H}_0 + g\mathbf{H}_1 \text{ where } [\mathbf{H}_0, \mathbf{H}_1] = 0.$$

   This entails that these Hamiltonians can be simultaneously diagonalized and so, the eigenfunctions are independent of g even though the eigenvalues vary with g. Then, there may be a level crossing where an excited level becomes the ground state at $g = g_c$, creating a point of non-analyticity of the ground state energy as a function of g.

- For an infinite lattice, the possibilities are richer. An avoided level-crossing between the ground and an excited state in a finite lattice could become progressively sharper as the lattice size increases, leading to a non-analyticity at $g = g_c$ in the infinite-lattice limit. Any point of non-analyticity in the ground state energy of the infinite lattice system as a quantum phase transtion. The non-analyticity could be either the limiting case of an avoided level crossing, or an actual level crossing, the first situation being the most common. The phase transition is usually accompanied by a qualitative change in the nature of the correlations in the ground state.

In particular, second order quantum phase transitions are transtions at which the characteristic energy scale of fluctuations above the ground state vanishes as $g \to g_c$. Let Δ be the scale characterizing some significant spectral density of fluctuations at zero temperature for $g \neq g_c$. Thus, Δ could be the energy of the lowest excitation above the ground state. If this is non-zero (ie. there is an energy gap Δ), or if there are excitations at arbitrarily low energies in the infinite lattice limit (ie. the energy spectrum is gapless), Δ is the scale at which there is a qualitative change in the nature of the frequency spectrum from its lowest frequency to its higher frequency behaviour. In most cases, it is the case that

$$\Delta_{\pm} \underset{g \to g_c}{\sim} J|g - g_c|^{zv}, \qquad \text{where } J \text{ is the energy scale of a characteristic microscopic coupling,}$$
$$\text{and where } zv \text{ is the critical phase exponent, which is universal.}$$

A critical phase exponent is independent of most of the microscopic details of the $\mathbf{H}(g)$-Hamiltonian. The previous behaviour holds both for $g > g_c$ and $g < g_c$ with the same value of the exponent zv, but with different non-universal constant of proportionality. In addition to a vanishing energy scale, second order quantum phase transitions invariably have a diverging characteristic length scale ζ, which could the length scale determining the exponential decay of equal time correlations in the ground state or the length scale at which some characteristic crossover occurs to the correlation at the longest distances. This length diverges as

$$\zeta^{-1} \sim \Lambda|g - g_c|^{v}, \qquad \begin{array}{l} \text{where } v \text{ is a critical exponent,} \\ \text{and where } \Lambda \text{ is an inverse length scale ("a momentum} \\ \text{cutoff") of order the inverse lattice spacing.} \end{array}$$

The ratio of the exponents for energy and length is z, the dynamic critical exponent: the characteristic energy scale vanishes as the z-th power of the characteristic inverse length scale

$$\Delta \sim \zeta^{-z}.$$

It is important to notice that the previous discussion concerns only the ground state of the system. Thus, quantum phase transitions occur only at zero temperature. Since all experiments are at necessarily non-zero temperature, a central task of the theory of quantum phase transitions is to describe the consequences of this $T = 0$-singularity on physical properties at $T > 0$. It turns out that working outward from the critical point $g = g_c$, and $T = 0$, is a powerful way of understanding and describing the thermodynamic properties of numerous systems over a broad range of values $|g - g_c|$ and T.

Quantum versus classical phase transitions. There are two important possibilities for the $T > 0$ phase diagram of a system near a quantum critical point,

- either the thermodynamic singularity is present only at zero temperature and all $T > 0$ properties are analytic as a function of g near $g = g_c$,
- or there is a curve of $T > 0$ second order phase transitions (this is a line at which the thermodynamic free energy is not analytic) which terminates at the $T = 0$-quantum critical point, $g = g_c$.

In particular, for the second kind of transitions, in the vicinity of such curve, the typical frequency at which the important long distance degrees of freedom fluctuate, ω_{typ}, satisfies

$$\hbar\omega_{\text{typ}} << k_B T.$$

Under

2. SPIN CHAIN MODELS

2.1. **Classical Spin Systems.** Classical spin systems are idealized versions of magnets. Although many magnetic phenomena in materials are inherently quantum mechanical, many properties are well described at least qualitatively by classical spin systems.

Each degrees of freedom of a spin system models a magnetic moment at a fixed location. A classical spin is simply the angular momentum vector of a spinning particle, and so can be represented by an arrow, which is assumed to have fixed length. For a spin system to be well-defined, the following requirements are needed

- constraints on where the arrow is allowed to point. For example, in the XY-model, the arrow is constrained to point in a two-diensional plane, and in the Ising model, the arrow is allowed to point in only directions. More complicated examples are the sigma-models, where instead of an arrow, the degrees of freedom take values on some manifold. In this language, the spins in the XY model take values on a circle.

- Where the spins physically are. For example, they may be at each point of some lattice in some dimension or they could be continuously distributed (ie. an arrow for every point in space eg. a field theory).

- How the spins themselves interact with one another ie. the energies associated with the possible spin configurations. Two major types of interaction are the ferromagnetic, where the energy is lowered when two spin are aligned, and antiferromagnetic, where the energy is lowered when they point in opposite directions. A given model can include interactions of both types.

Many many-body physical systems can effectively be treated as being on a lattice. For example, many of these systems are often well treated by utilizing a tight-binding model, where each electron is treated as being located at a fixed nucleus and so live at particular points in space. In some situations, the physics itself arises from the interplay between the degrees of freedom and the particular lattice they live on eg. geometrical frustration in the two-dimensional antiferromagnetic Ising model on a triangular lattice. Calling the two allowed directions in the Ising model as "up" and "down", antiferromagnetic interactions make adjacent spins different. On the square lattice, it is possible to make a low-energy state by alternating up and down spins, but on a triangular lattice this is not the case. Around a triangle, there must be at least two mutually up or down spins adjacent to each other. Such a bond in said to be unsatisfied and so the spins are frustrated, yielding completely new properties different

fro the unfrustrated model.

2.1.1. *Formalization of Statistical Mechanics.* A formal definition for a classical general spin chain, as a dynamical system, requires the use of measure theory terminology.

Some technical aspects. <u>**An introduction to Measure Theory**</u>
A **measure** is a mathematical device which reflects the notion of quantity for a given set. Let $\mathbf{X}$ be a set, then each subset $\mathbf{U} \in \mathbf{X}$ is assigned a positive real number $\mu[\mathbf{U}]$. Thus, the measure is a function

$$\mu : \mathrm{Dom}(\mu) \subset 2^{\mathbf{X}} \to \mathbb{R},$$

where $2^{\mathbf{X}} = \{\mathbf{S} \in \mathbf{X}\}$ is the power set of $\mathbf{X}$. However, it's usually impossible to define a satisfactory notion of quantity for all subsets of $\mathbf{X}$. Therefore, it is useful to consider $\mathrm{Dom}(\mu) \in \mathbf{X}$ ie. only some subsets of $\mathbf{X}$ will be measurable. Before defining a proper measure, it's domain must be specified first. This domain will be a collection of subsets of the space $\mathbf{X}$, called a σ-**algebra**.

Let $\mathbf{X}$ be a nonempty set, which in the following sections will be a sample space. A σ-**algebra** over $\mathbf{X}$ is a collection $\mathcal{F} = \{\mathbf{U}_i\}_{i \in \mathbb{N}}$ of subsets of $\mathbf{X}$ with the following properties:

- $\mathcal{F}$ contains the set $\mathbf{X}$: $\mathbf{X} \subset \mathcal{F}$.

- $\mathcal{F}$ is closed under <u>complementation</u>: if $\mathbf{U} \subset \mathcal{F}$, then $\mathbf{U}^c = (\mathbf{X}/\mathbf{U}) \subset \mathcal{F}$.

- $\mathcal{F}$ is closed under <u>countable unions</u> ie. $\{\mathbf{U}_i\}_{i \in \mathbb{N}} \mid \mathbf{U}_i \subset \mathcal{F} \Rightarrow \underset{i \in N}{\cup} \mathbf{U}_i \subset \mathcal{F}$.

- $\mathcal{F}$ is closed under <u>countable intersections</u> ie. $\{\mathbf{U}_i\}_{i \in \mathbb{N}} \mid \mathbf{U}_i \subset \mathcal{F} \Rightarrow \underset{i \in N}{\cap} \mathbf{U}_i \subset \mathcal{F}$.

Technically, the fourth property, is a corollary and follows directly from the second and third properties and due to De Morgan's laws.

A **measurable space** is an ordered pair $(\mathbf{X}, \mathcal{F})$, where $\mathbf{X}$ is a set and where $\mathcal{F}$ is a σ-algebra on $\mathbf{X}$. Some of the most common examples of σ-algebras are

- Two examples of trivial σ-algebras are:
  - for any set $\mathbf{X}$ the collection $\{\emptyset, \mathbf{X}\}$ is a σ-algebra. This first example is far too small to be of any use.
  - The power set $2^{\mathbf{X}}$ is also a σ-algebra. Note that for large sets, the power set becomes too large to be manageable.

- A more manageable σ-algebra is the one induced by the (co-)countable sets. Let $\mathcal{M}$ be the most conservative collection of "manageable" sets, this is

$$\mathcal{M} = \Big\{ \{x\} : x \in \mathbf{X} \Big\},$$

ie. the set of all the singleton subsets of $\mathbf{X}$. Then $\mathcal{C} = \sigma(\mathcal{M}))$ is the σ-algebra of **countable and co-countable sets**

$$\mathcal{C} = \{\mathbf{C} \subset \mathbf{X} \mid \text{either } \mathbf{C} \text{ is countable, or } \mathbf{X}/\mathbf{C} \text{ is countable}\}.$$

If $\mathbf{X}$ is itself finite or countable, then $\mathcal{C} = \mathcal{P}(\mathbf{X})$.

- Another example are the partition algebras. Let $\mathbf{X}$ be a set. Then a **partition** of $\mathbf{X}$ is a collection $\mathcal{P} = \{\mathbf{P}_i\}_{i=1}^N$ of disjoint subsets, such that $\mathbf{X} = \sqcup_{n=1}^N \mathbf{P}_N$. These subsets $\mathbf{P}_i$ are called the **atoms** of the partition. Then, the σ-algebra generated by $\mathcal{P}$ is the collection of all possible unions of $\mathcal{P}$-atoms:

$$\sigma(\mathcal{P}) = \{\sqcup_{j=1}^{k} \mathbf{P}_{n_j} \mid \{n_j\}_{j=1}^{k} \in \mathbb{N}_{[1,\cdots,N]}\}.$$

Therefore if $\text{card}[\mathbf{P}] = N$, then $\text{card}[\sigma(\mathbf{P})] = 2^N$.

If $\mathcal{Q}$ is another partition, we say $\mathcal{Q}$ redefines $\mathcal{P}$ ($\mathcal{Q} \prec \mathcal{P}$) if, for every $\mathbf{P} \in \mathcal{Q}$, there are $\{\mathbf{Q}_i\}_{i=1}^{N} \in \mathcal{Q}$ so that $\mathbf{P} = \sqcup_{j=1}^{N} \mathbf{Q}_j$. In said case we have

$$\mathcal{P} \prec \mathcal{Q} \Leftrightarrow \sigma(\mathcal{P}) \subset \sigma(\mathcal{Q}).$$

- **Borel σ-algebra of $\mathbb{R}$:**

  Let $\mathbf{X} = \mathbb{R}$ be the real numbers and let $\mathcal{M}$ be the set of all open intervals in $\mathbb{R}$:

  $$\mathcal{M} = \{(a,b) : -\infty \leq a < b \leq \infty\},$$

  then the σ-algebra $\mathcal{B} = \sigma(\mathcal{M})$ contains all open subsets of $\mathbb{R}$, all closed subset, all countable intersections of open subsets, countable unions of closed subsets, etc. For example, $\mathcal{B}$ contains, as elements, the set $\mathbb{Z}$ of integers, the set $\mathbb{Q}$ of rationals and the set $\mathbb{I}$ of irrationals. Then, $\mathcal{B}$ is called the **Borel σ-algebra** of $\mathbb{R}$.

In general, let $\mathbf{X}$ be a topological space and let $\mathcal{M}$ be the set of all open subsets of $\mathbf{X}$. The σ-algebra $\sigma(\mathcal{M})$ is the **Borel σ-algebra** of $\mathbf{X}$ and is denoted by $\mathcal{B}(\mathbf{X})$. It contains all open sets and closed subsets of $\mathbf{X}$, all countable intersections of open sets (called $G\delta$ sets), all countable unions of closed sets (called $F\sigma$ sets). For example, if $\mathbf{X}$ is Hausdorff, then $\mathcal{B}(\mathbf{X})$ contains all countable and co-countable sets.

Let $(\mathbf{X}, \mathcal{F})$ be a measurable space. A **measure** on $\mathcal{F}$ is a map $\mu : \mathcal{F} \to \mathbb{R}_+$, which is **countably additive** ie.

$$\text{If } \{\mathbf{Y}_i\}_{i=1} \mid \mathbf{Y}_i \in \mathcal{F} \to \mu\Big[\sqcup_{n=1}^{\infty} \mathbf{Y}_n\Big] = \sum_{n=1}^{\infty} \mu[\mathbf{Y}_n].$$

Then a **measure space** is an ordered triple $(\mathbf{X}, \mathcal{F}, \mu)$, where $\mathbf{X}$ is a set, $\mathcal{F}$ is a σ-algebra and μ is a measure on $\mathcal{F}$. Thus, μ assigns a size to the $\mathcal{F}$-measurable subsets of $\mathbf{X}$. Some important measures and measureable spaces are

- **The counting measure** assigns, to any set, the cardinality of that set.

  $$\mu[\mathbf{S}] = \text{card}[\mathbf{S}].$$

  This measure provides no means of distinguishing between sets of the same cardinality, only being useful in finite measure spaces.

- A **finite measure space** is made up by a finite set $\mathbf{X}$, and a σ-algebra $\mathcal{F} = \mathbf{X}$. Then a measure μ on $\mathbf{X}$ is entirely defined by some function $f : \mathbf{X} \to \mathbb{R}_+$. For any subset $\{x_i\}_{i=1}^{N}$ we then define

  $$\mu\Big(\{x_i\}_{i=1}^{N}\Big) = \sum_{i=1}^{N} f(x_i).$$

  Discrete probability theory is based on at most countable sample spaces Ω, where all subsets of Ω may be treated as events, and thus elements of the σ-algebra $\mathcal{F}$.

- **Discrete measures**: If, instead, the sample space is uncountable, (eg. it has a discrete part a non-discrete part), it can be decomposed into **atoms**. The atoms are an at most countable (maybe even empty) set, whose probability is the sum of probabilities of all atoms. If this sum is equal to 1, then all other points can safely be excluded from the sample space, returning to the discrete case. Otherwise, if the sum of probabilities of all atoms is between 0 and 1, then the measure space decomposes into a discrete atomic part and a non-atomic part. If $(\mathbf{X}, \mathcal{F}, \mu)$ is a measure space, then an **atom** of μ is a subset $\mathbf{A} \in \mathcal{F}$ such that

– it has positive measure $\mu[\mathbf{A}] = A > 0$.

– and it contains no set of smaller positive measure, ie. $\forall \mathbf{B} \subset \mathbf{A}$, either $\mu[\mathbf{B}] = A$ or $\mu[\mathbf{B}] = 0$.

For example, in the finite measure space above, the singleton set $\{x_n\}$ is an atom if $f(x_n) > 0$. The measure space $(\mathbf{X}, \mathcal{F}, \mu)$ is called discrete if it can be written as

$$\mathbf{X} = \mathbf{Z} \sqcup (\sqcup_{n=1}^\infty A_n),$$

where $\mu[\mathbf{Z}] = 0$ and where $\{A_n\}_{n=1}^\infty$ is a collection of atoms. Note that any finite measure space is discrete.

For example, consider the set $\mathbf{X} = \{1, 2, ..., 9, 10\}$ and let the σ-algebra Σ be the power set of $\mathbf{X}$. Define the measure μ of a set to be its cardinality, that is, the number of elements in the set. Then, each of the singletons i, for i = 1, 2, ..., 9, 10 is an atom. On the other hand, consider the Lebesgue measure on the real line. This measure has no atoms.

- **Density Functions**: Let $\rho : \mathbb{R}^n \to \mathbb{R}$ be a positive, integrable function on $\mathbb{R}^n$. For any $\mathbf{B} \in \mathcal{B}(\mathbb{R}^n)$, then a measure can be defined as

$$\mu_\rho(\mathbf{B}) = \int_{\mathbf{B}} \rho,$$

where ρ is the **density function** for μ.

Then, the ordered triple $(\mathbf{X}, \mathcal{F}, \mu)$ is a **probability space** if μ is a **probability measure**. This is, a measure μ on $\mathbf{X}$ is a **probability measure** if $\mu[\mathbf{X}] = 1$ [1].

---

[1]**Stochastic processes** are a particular kind of probability measures, which represent a system randomly evolving in time. Let $\mathcal{S}$ be any randomly-evolving complex system, let $\mathbf{X}$ be the set of all possible states of the system $\mathcal{S}$ and let $\mathbb{T}$ be a set representing time. For example,

- If $\mathcal{S}$ is a rolling die, then $\mathbf{X} = \{1, 2, 3, 4, 5, 6\}$ and $\mathbb{T} = \mathbb{N}$ indexes the successive dice rolls.

- If $\mathcal{S}$ is a publically traded stock, then its state is its price. Thus $\mathbf{X} = \mathbb{R}$. If we assume trading occurs continuously when the market is open, and let each trading day have length $c < 1$, then one representation of market time is $\mathbb{T} = \sqcup_{n=1}^\infty [n, n+c]$.

- If $\mathcal{S}$ is a weather system, then its state can be represented by a large array of data $\mathbf{x} = [x_1, \cdots, x_n]$. Thus $\mathbf{X} = \mathbb{R}^n$ and since the weather evolves continuously $\mathbb{T} = \mathbb{R}$.

The random evolution of $\mathcal{S}$ is represented by assigning a probability to every possible *history*. A history is an assignment of a state in $\mathbf{X}$ to every moment in time (ie. in $\mathbb{T}$). In other words, it's a function $f : \mathbb{T} \to \mathbf{X}$. The set of all possibly histories is $\mathbf{H} = \mathbf{X}^{\mathbb{T}}$ ie. the set of all functions $f : \mathbb{T} \to \mathbf{X}$. The σ-algebra on $\mathbf{H}$ is usually a **cylinder algebra**.

- Let $(\mathbf{X}_\lambda, \mathcal{X}_\lambda)$ be measurable spaces for all $\lambda \in \Lambda$, where Λ is some (possibly uncountably infinite) indexing set. Consider the cartesian product $\underset{\lambda \in \Lambda}{\times} \mathbf{X}_\lambda$. Let

$$\mathcal{M} = \left\{ \underset{\lambda \in \Lambda}{\times} \mathbf{U}_\lambda \mid \forall \lambda \in \Lambda \text{ and } \mathbf{U}_\lambda \in \mathcal{F}_\lambda \text{ and } \mathbf{U}_\lambda = \mathbf{X}_\lambda \text{ for all but finitely many } \lambda \right\},$$

such subsets are called **cylinder sets** in $\mathbf{X}$ and $\sigma(\mathcal{M})$ is the **cylinder σ-algebra** denoted by $\underset{\lambda \in \Lambda}{\times} \mathbf{X}_\lambda \mathcal{F}_\lambda$. If the $\mathbf{X}_\lambda$ are topological spaces with Borel σ-algebras $\mathcal{F}_\lambda$, and we endow $\mathbf{X}$ with the Tychonoff product topology, then $\underset{\lambda \in \Lambda}{\times} \mathcal{F}_\lambda$ is the Borel σ-algebra of $\mathbf{X}$.

Suppose that $\mathbf{X}$ has a σ-algebra $\mathcal{F}$, then it follows $\mathbf{H}$'s σ-algebra is $\mathcal{H} = \underset{t \in \mathbb{T}}{\times} \mathcal{F}_t$. An **event** is an element of $\mathcal{F}$ and thus corresponds to a cylinder set, a countable union of cylinder sets etc. Suppose, for all $t \in \mathbb{T}$, that $\mathbf{U}_t \in \mathcal{F}$ with $\mathbf{U}_t = X$ for all but finitely many t. The cylinder set $\mathbf{U} = \prod_{t \in \mathbb{T}} \mathbf{U}_t$ thus corresponds to the assertion: "for every $t \in \mathbb{T}$, at time t, the state of $\mathcal{S}$ was inside $\mathbf{U}_t$. A probability measure on $(\mathbf{H}, \mathcal{H})$ is then a way of assigning probabilities to such assertions.

Gibbs Random Fields. A **Gibbs random field** is defined on a lattice, as follows

- A lattice, which is a countable set $\mathbb{L}$ and a set $\mathcal{L}$, the set of all finite subsets of $\mathbb{L}$.
- A single-spin space, which is a probability space $(S, \boldsymbol{\mathcal{S}}, \lambda)$, where S is the set of all possible spin-values and is assumed to be discrete, $\boldsymbol{\mathcal{S}}$ is the σ-algebra associated with this set. By definition, $\boldsymbol{\mathcal{S}} \subseteq 2^S$, where 2^S is the powerset of S, and where λ is the probability measure on this single-spin space.
- A configuration space $(\Omega, \mathcal{F})$, where $\Omega = S^{\mathbb{L}}$ is the event space and where $\mathcal{F} = \mathcal{S}^{\mathbb{L}}$.
- Given a configuration $\omega \in \Omega$ and a subset $\Lambda \subset \mathbb{L}$, the restriction of ω to Λ is

$$\omega_\Lambda = (\omega(t))_{t \in \Lambda}.$$

In particular, if $A_1 \cap A_2 = \emptyset$ and $A_1 \cup A_2 = \mathbb{L}$, then the configuration $\omega_{\Lambda_1}\omega_{\Lambda_2}$ is the configuration whose restrictions to Λ_1 and Λ_2 are ω_{Λ_1} and ω_{Λ_2}, respectively.

- For each subset $\Lambda \subset \mathbb{L}$, $\mathcal{F}_\Lambda$ is the σ-algebra generated by the family of functions

$$(\sigma(t))_{t \in \Lambda}, \text{ where } \sigma(t)(\omega) = \omega(t).$$

The union of these σ-algebras as Λ varies over $\mathcal{L}$ is the algebra of cylinder sets on the lattice.
- The potential is a family of functions

$$\Phi = (\Phi_A)_{A \in \mathcal{L}} \text{ where } \Phi_A : \Omega \to \mathbb{R},$$

such that
1) $\forall A \in \mathcal{L}$, Φ_A is $\mathcal{F}_\Lambda$-measurable, meaning it depends only on the restriction ω_A, and does so measurably.
2) $\forall A \in \mathcal{L}$, $\forall \omega \in \Omega$, then

$$\exists H_\Lambda^\Phi(\omega) = \sum_{\substack{A \in \mathcal{L} \\ A \cap \Lambda \neq \emptyset}} \Phi_A(\omega),$$

where Φ_A is interpreted as the contribution to the total energy, ie. the Hamiltonian, associated to the interaction among all the points of finite set A. Then, $H_\Lambda^\Phi(\omega)$ is the contribution to the total energy of all the finite sets A that meet Λ. Typically, the total energy may be infinite.

3) The Hamiltonian in $\Lambda \in \mathcal{L}$ with boundary conditions $\bar{\omega}$, for the potential Φ, is defined by

$$H_\Lambda^\Phi(\omega|\bar{\omega}) = H_\Lambda^\Phi(\omega_\Lambda \bar{\omega}_{\Lambda^c}) \text{ where } \Lambda^c = \mathbb{L}/\Lambda.$$

- The partition function in $\Lambda \in \mathbb{L}$ with boundary conditions $\bar{\omega}$ and inverse temperature $\beta > 0$, for the Φ-potential and λ previously defined, is given by

$$\mathcal{Z}_\Lambda^\Phi(\bar{\omega}) = \int \lambda^\Lambda(d\omega) exp(-\beta H_\Lambda^\Phi(\omega|\bar{\omega}), \text{ where } \lambda^\Lambda(d\omega) = \prod_{t \in \Lambda} \lambda(d\omega(t)) \text{ is the product measure.}$$

As dynamical systems, the Ising model or lattice gas can be obtained by first ignoring the momenta and assuming that the particles' positions are restricted to a discrete subset of $\mathbb{R}^d$. In general, this subset is taken to be the d-**dimensional cubic lattice**. Then,

(1) $\qquad$ the lattice is simply $\mathbb{L} = \mathbb{Z}^d$, where $\mathbb{Z}^d = \{\mathbf{i} = (i_1, \cdots, i_d) \in \mathbb{R}^d \mid i_k \in \mathbb{Z} \; \forall k \in \mathbb{N}_{[1,d]}\}$.

In other words, the system is imagined as living on $\mathbb{R}^d$ and said space is composed of small cells, so that each cell can accommodate at most one particle. On each one of these cells, a single-spin space is then defined as $S_i = \{-1, 1\}^i$, $i \in \mathbb{Z}$. To describe the model's microstates, consider a finite region $\Lambda \subset \mathbb{Z}^d$, representing the vessel and an **occupation number** n_i defined as

(2) $\qquad n_i : \Lambda \to \{0,1\} \Rightarrow$ then the set of microstates is $\Omega_\Lambda = \{0,1\}^{\times\Lambda}$.

The model automatically includes a short-range repulsion between the particles since no two particles are allowed to share a same cell. The attractive part of the interaction can then be included into the Hamiltonian:

$$\forall \mathbf{n} = (n_i)_{i\in\Lambda}, n_i \in \Omega_\Lambda, \ \mathbf{H}(\mathbf{n}) = \sum_{\{i,j\}\subset\Lambda} \mathcal{J}(i-j)n_i n_j,$$

which is completely similar to a classical, continuous, Hamiltonian $\mathbf{H}(\mathbf{q})$ and where the function $\mathcal{J} : \mathbb{Z}^d \to \mathbb{R}$ depends, *a priori*, only on the distance between the i-th and j-th cells. Note that the contribution of a pair of cells $\{i,j\}$ is zero if they do not both contain a particle. The number of particles in Λ is given by

$$N_\Lambda(\mathbf{n}) = \sum_{i\in\Lambda} n_i.$$

Thus the basic object classical statistical mechanics is the Boltzmann weight $\mathbb{P}(n)$. In thermal equilibrium, this is a probability measure that gives the probability that a system will be in a certain configuration $\mathbf{n} \in \Omega_\Lambda$, in terms of said configuration's energy and system's temperature. In other words,

(3) $$\mathbb{P}_{\Lambda,\beta,N} : \mathbf{n} \in \Omega_\Lambda \to \mathbb{R}_{[0,1]} \text{ where } \mathbb{P}_{\Lambda,\beta,N}(\mathbf{n}) = \frac{e^{-\mathbf{H}(\mathbf{n})}}{\mathcal{Z}_{\Lambda,\beta,N}},$$

where $\mathcal{Z}_{\Lambda,\beta,N}$ is the partition function defined on the finite region $\Lambda \subset \mathbb{Z}^d$, at inverse temperature β with N particles. The Boltzmann weight is defined by the requirement that the probabilities sum to one. Then, the two relevant ensemble partition functions can be defined, as follows

$$\mathcal{Q}_{\Lambda,\beta,N} = \sum_{\eta\in\Omega_{\Lambda,N}} e^{-\beta\mathbf{H}(\mathbf{n})} \qquad\qquad \Theta_{\Lambda,\beta,N} = \sum_N e^{\beta\mu N} \sum_{\eta\in\Omega_{\Lambda,N}} e^{-\beta\mathbf{H}(\mathbf{n})}$$

$$\text{where } \Omega_{\Lambda,N} = \{n \in \Omega_\Lambda | N_\Lambda(n) = N\}$$

Note that if the degrees of freedom take on continuous values, or the model is defined in the continuum, then the sums are replaced by integrals. Note that, indeed, the probability of a given configuration increases as the energy gets lower, and conversely, that as the temperature gets higher and higher, the energies involved must get larger and larger to make a difference. Note as well that if all the energies are shifted by some constant E_0, the probabilities are left unchanged, since both the numerator and denominator are multiplied by the same constant, namely $e^{-\beta E_0}$.

Consider then the Ising model, with nearest-neighbour interactions (coupling constant J) and a magnetic field (h). Its regular crystalline structure, corresponding to the positions of the magnet's atoms by a finite, non-oriented graph $G = (\Lambda, \mathcal{E})$, whose set of vertices is $\Lambda \subset \mathbb{Z}^d$. Then, the **box of radius** n is

$$B(n) = \{-n,\cdots,n\}^d$$

The edges of the graph will most often be between nearest neighbours, that is, pairs of vertices i, j such that $||j-i||_1 = 1$ [2]. Two nearest-neighbours spins are denoted by $i \sim j$. Then, the set of edges in the box $B(n)$ is $\{\{i,j\} \subset B(n) : i \sim j\}$.

The Ising model is defined by first assuming that a spin is located at each vertex of a graph $G = (\Lambda, \mathcal{E})$, such that each single-spin space has cardinality two, ie. each spin can only take two values. It follows, then, that to describe a microstate, a variable σ_i taking two possible values (± 1) is associated to each vertex $i \in \Lambda$, which is called the **spin** at i. A microstate of the system is called a **configuration**, and is thus an element

$$\sigma \in \Omega_\Lambda = \{0,1\}^{\times\Lambda}.$$

---

[2] In this discrete setting, the norm is defined such that $||i||_1 = \sum_{k=1}^d |i_k|$

An N-site chain, with finite N, has a configuration space, which is a discrete space, which is a 0-manifold. This 0-manifold has zero topological dimension and 2^N cardinality. The microscopic interactions among the spins are defined such that

1) There are only interactions between pairs of sins located at neighbouring vertices. That is, it is assumed that the spins at two distinct vertices $i, j \in \Lambda$ interact if and only if the pair $\{i, j\}$ is an edge of the graph.

2) The interaction favours agreement of spin values. In the simplest form of the Ising model, this is done in the simplest possible way: a pair of spins at the endpoints i and j of an edge, decreases the overall configuration's energy if they agree $\sigma_i = \sigma_j$ and increases if they differ. More precisely, the spins at the endpoints of the $\{i, j\}$−edge contributes to the total energy by an amount $-\sigma_i \sigma_j$. Therefore, configurations in which most pairs of neighbours are aligned to have smaller energy.

3) Spins align with the external magnetic field. Assume that a constant external magnetic field of $h \in \mathbb{R}$ intensity, oriented along the same direction as the spins, acts on the system. Its interaction with the i-th spin contributes an $-h\sigma_i$ to the total energy. That is, when the magnetic field is positive, the configurations with most of their spins equal to +1 have smaller energy.

Then, the energy of a given configuration σ is obtained by summing the interactions over all pairs and by adding the interaction of each spin with the external magnetic field, thus yielding the **Ising Hamiltonian**

$$(4) \qquad \mathbf{H}_{\Lambda;h}(\sigma) = -J \sum_{i,j \in \Lambda i \sim j} \sigma_i \sigma_j - h \sum_{i \in \Lambda} \sigma_i, \text{ with } \sigma \in \Omega_\Lambda.$$

Since it favours local alignment of the spins, the Hamiltonian of the model is said to be ferromagnetic, which is not to say it behaves like a ferromagnet. The Gibbs measure/distribution is then denoted by

$$(5) \qquad \mathbb{P}_{\Lambda,\beta,h} : \mathbf{n} \in \Omega_\Lambda \to \mathbb{R}_{[0,1]} \text{ where } \mathbb{P}_{\Lambda,\beta,h}(\sigma) = \frac{e^{-\beta \mathbf{H}_{\Lambda;h}(\sigma)}}{\mathcal{Z}_{\Lambda,\beta,h}}.$$

Note that, in the absence of a magnetic field, even though local spin alignment is favoured by the Hamiltonian, neither of the orientations (+1 or -1) is globally favoured. Namely, if $-\sigma$ denotes the spin-flipped configuration in which $(-\sigma)_i = -\sigma_i$, then $\mathbf{H}_{\Lambda;0}(\sigma) = \mathbf{H}_{\Lambda;0}(-\sigma)$, this implies

$$\mathbb{P}_{\Lambda,\beta,0}(-\sigma) = \mathbb{P}_{\Lambda,\beta,0}.$$

The model is then said to be invariant under global spin flip. When $h \neq 0$, this symmetry no longer holds. Then the expectation of an observable $f : \Omega_\Lambda \to \mathbb{R}$, under $\mu_{\Lambda,\beta,h}$, is denoted by $\langle f \rangle_{\Lambda,\beta,h}$. The parameter J is the coupling strength, if $J > 0$ this models describes a ferromagnetic coupling, while $J < 0$ describes an antiferromagnetic coupling. From the partition function, the expectation values of physical quantities can be calculated. For example

- Internal energy $\langle E \rangle_{\Lambda,\beta,h} = \frac{1}{\mathcal{Z}} \sum_n E_n e^{-\beta E_n} = -\frac{\partial}{\partial \beta} \log \mathcal{Z}$.

- Specific heat $C = \frac{\partial}{\partial T} \langle E \rangle_{\Lambda,\beta,h} = k \frac{\partial}{\partial T} T^2 \frac{\partial \log \mathcal{Z}}{\partial T}$

- Two-point correlator in the Ising model,

$$\langle \sigma_a \sigma_b \rangle_{\Lambda,\beta,h} = \sum_{\sigma_i} \sigma_a \sigma_b \mathbb{P}_{\Lambda,\beta,h}(\sigma_i) = \frac{1}{\mathcal{Z}} \sum_{\sigma_i = \pm 1} \sigma_a \sigma_b e^{-\beta \mathbf{H}_{\Lambda;h}(\sigma)},$$

which describes how the degrees of freedom at one point are affected by the degrees of freedom at another point. If $\langle \sigma_a \sigma_b \rangle = 1$, the two spins are aligned, while if $\langle \sigma_a \sigma_b \rangle = -1$ the two spins are anti-aligned, and if $\langle \sigma_a \sigma_b \rangle = 0$, they are uncorrelated.

Three kinds of phases. Even in simple interacting systems, the two point function will not have such simple values except in some extreme limits. However, knowing its dependence on the separation is valuable information, and it implies that the behaviour of the correlators is one of the best ways of understanding and characterising different phases of matter. Consider a correlator defined so that in the non-interacting limit, ie. the $T \to \infty$, it vanishes. Letting L be the system size, there are three main types of behaviour of such a two-point function as $L >> |a - b| >> 1$, as follows

- **ordered:**

$$\langle \sigma_a \sigma_b \rangle \sim C \neq 0,$$

In an ordered spin system, the spins are correlated arbitrarily far from each other: if a spin, at a given position, points in some direction, then another spin is going to most likely point in the same direction. For example, in an antiferromagnet, order typically means that spins at an odd number of sites apart, are antialigned, while those an ever number of sites apart are aligned.

Another way to characterize order is in terms of a one-point function, of a given variable whose thermodynamic average vanishes in the non-interacting limit. In the Ising model, this is simply $\langle \sigma_a \rangle$, which is the probability that a given spin is either up or down. In principle, this is easier to compute than the two-point function but, in many cases including the Ising model, it vanishes for all values of a, b, J. Indeed, this follows from the Ising model's $\mathbb{Z}_2$-symmetry $\sigma_i \to -\sigma_i$. Then, $E(\{\sigma_i\}_{i=1}^N) = E(-\{\sigma_i\}_{i=1}^N)$. Therefore, as long as the boundary conditions remain invariant under spin flip, it holds that

(6)
$$-\langle \sigma_a \rangle = -\sum_{\sigma_i = \pm 1} \sigma_i e^{-\beta E(\{\sigma_i\})} = \sum_{\sigma_i = \mp 1} s_a e^{-\beta E(\{s_i\})} \text{ where } s_i = -\sigma_i.$$

However, since the sum is performed over a finite number of lattice sites, the s_i-variables on the last equality may be renamed as σ_i. Thus, the last expression is simply $\langle \sigma_a \rangle = 0$.

- **disordered:**

$$\langle \sigma_a \sigma_b \rangle \sim e^{-\frac{|a-b|}{\epsilon}},$$

A more general way of characterizing order is to use the previous definition, ie. a non-vanishing value of a two-point function at arbitrarily large separation. Then, the order parameter can be defined as the square root of this absolute value.

- **critical:**

$$\langle \sigma_a \sigma_b \rangle \sim \frac{1}{|a - b|^{2x}}.$$

In a disordered phase, the two-point function class off to zero exponentially fast. Thus a conveniet measure is the correlation length ϵ. A useful heuristic argument is to think of the degrees of freedom closer than ϵ as being correlated, and those further apart as being uncorrelated. One thing to bear in mind is that not all disordered systems behave in the same way. Some systems are disordered by definition, but possess a topological order, where an expectation value of some non-local quantities has a non-vanishing expectation value.

The system's criticality is the most difficult and physically meaningful case. The correlator is said to be algebraically decaying. In particular, in the algebraic decay, the spatial dependence does not involve in any way a dimensional parameter, but rather only a dimensionless parameter x. Thus, this behaviour is characteristic of a critical phase, where the system is scale-invariant, given that there is no parameter like a correlation length. The reason for the factor of 2 in the x-quantity's definition stems from this picture, in the continuum limit, x can be thought of as a "scaling dimension" of the spin field.

Note that, this picture is oversimplified since different types of correlators can occur in the same system.

A common case in statistical mechanics occurs when, at temperatures much larger than any parameter with dimensions of energy - like J in the Ising model-, one can effectively neglect the energy altogether. The partition function is then a sum over all allowed configurations with the same weight each. Then, the two-point function of any local operators vanishes quickly as they are brought apart (as long as non-local constraints are not present). The system is then in a disordered phase at sufficiently large temperatures. At low enough temperatures, the opposite occurs: since the Boltzmann weight is proportional to $e^{-E/T}$, the very low-energy configurations are much more likely. This can lead to order, the lowest-energy configurations have some regular pattern and so there is a non-vanishing local order parameter. When there is a disordered phase at high-temperature and an ordered phase at high temperature as well, there must be a phase transition in between. There are two categories of phase transitions, often referred to as first-order and continuous. In the latter case, correlators become algebraically decaying, and then the system is critical. In the former case, the system remains disordered at the phase transition and abruptly changes its behaviour.

2.1.2. *Exact solution to the Ising Models.*

Zero-dimensional Ising model. A system with a finite number of spins, no matter in which directions, is commonly referred to a zero-dimensional model. In the $d = 0$-Ising model, there are just two degrees of freedom, called s_1 and s_2, both of which can only take two values $s_i \in S = \{\pm 1\}$ and where the configuration space is given by $(\Omega_\Lambda, \mathcal{F})$, where $\Omega_\Lambda = \{-1, 1\}^{\times 2}$ and where the σ-algebra $\mathcal{F} = 2^{\Omega_\Lambda}$. To compute the partition function, $\mathcal{Z}_{\Lambda,\beta,B}$, the system's energy must be known for every configuration $(s_1, s_2) \in \Omega_\Lambda$, which is assumed to be

$$(7) \qquad E(\mathbf{s}) = -Js_1s_2 - B(s_1 + s_2).$$

This energy models a system made of magnetic moments. Although, in general, the moments can point in any direction, in the Ising model they point up or down one axis - say the z-axis - along which an external magnetic field B is applied. The constant J has units of energy and B has units of magnetic moment. If $J > 0$, the model is ferromagnetic and the energy favors aligned spins while if $J < 0$ it models an antiferromagnetic system, which favours anti-parallel spins. In both cases, $B > 0$ is chosen so that the spins tend to align with the external field.

The partition function may be evaluated as

$$(8)$$

$$\mathcal{Z}_{\Lambda,K,h} = \sum_{\mathbf{s} \in \Omega_\Lambda} e^{Ks_1s_2 + h(s_1+s_2)} \text{ where } \begin{array}{l} K = \beta J \\ h = \beta B. \end{array} \Rightarrow \mathcal{Z}_{\Lambda,K,h} = \sum_{m=-\frac{1}{2}}^{\frac{1}{2}} e^{\beta hm} 2\cosh\left(\frac{h+Km}{2}\right) = 2\cosh(2h)e^K + 2e^{-K}.$$

At very low temperatures (very low T or very high β), the state of lowest energy will dominate the sum and thus, the spins are expected to be aligned with B and also with each other. At very high temperatures and vanishing β, all four states get equal weight and the spins will fluctuate independently of each other and the applied field. In effect, let the system's (spin-) average magnetization be $M = \frac{1}{2}(s_1 + s_2)$, in any given configuration[3]. The thermal average can then be found to be

$$(9)$$
$$\begin{aligned} \langle M \rangle_{\Lambda,K,h} &= \frac{1}{\mathcal{Z}_{\Lambda,\beta,h}} \sum_{\mathbf{s} \in \Omega_\Lambda} \frac{1}{2}(s_1 + s_2)e^{Ks_1s_2 + h(s_1+s_2)} = \frac{1}{2\mathcal{Z}_{\Lambda,K,h}} \frac{\partial \mathcal{Z}_{\Lambda,K,h}}{\partial h} \\ &= \frac{1}{2}\frac{\partial \log \mathcal{Z}_{\Lambda,K,h}}{\partial h}. \end{aligned}$$

In terms of the free energy $F(K, h)$, defined as $\mathcal{Z}_{\Lambda,K,h} = e^{-\beta F}$, where $-\beta F = \log\left(2\cosh(2h)e^K + 2e^{-K}\right)$ it yields

---

[3]Note that this is the average over all spins in the system, and not the thermal average

$$(10) \qquad \langle M \rangle_{\Lambda,K,h} = \frac{\partial[-\beta F(K,h)]}{\partial h} = \frac{\sinh 2h}{\cosh 2h + e^{-2K}} \text{ which, as expected } \begin{array}{l} \langle M \rangle \underset{h,K\to\infty}{\to} 1 \\ \langle M \rangle \underset{h,K\to 0}{\to} h. \end{array}$$

The thermal average of a particular spin can also be calculated, provided a soruce term of the form $h_1 s_1 + h_2 s_2$ is added in the Boltzmann weight, which couples the i th spin to its own independent field h_i. Then,

$$(11) \qquad \mathcal{Z}_{\Lambda,K,h} = \sum_{\mathbf{s}\in\Omega_\Lambda} e^{Ks_1 s_2 + \mathbf{h}\cdot\mathbf{s}} = e^{-\beta F(K,\mathbf{h})}$$

$$(12) \qquad \Rightarrow \langle s_i \rangle = \frac{\partial[-\beta F(K,\mathbf{h})]}{\partial h_i} = \frac{\partial \log \mathcal{Z}_{\Lambda,K,h}}{\partial h_i} = \frac{1}{\mathcal{Z}_{\Lambda,K,h}} \frac{\partial \mathcal{Z}_{\Lambda,K,h}}{\partial h_i}.$$

Taking the mixed derivative with respect to both h_1 and h_2 yields

$$
\begin{aligned}
\frac{\partial^2[-\beta F(K,\mathbf{h})]}{\partial h_1 \partial h_2} &= \frac{\partial}{\partial h_1}\left(\frac{1}{\mathcal{Z}_{\Lambda,K,h}}\frac{\partial \mathcal{Z}_{\Lambda,K,h}}{\partial h_2}\right) \\
&= \frac{1}{\mathcal{Z}_{\Lambda,K,h}}\frac{\partial^2 \mathcal{Z}_{\Lambda,K,h}}{\partial h_1 \partial h_2} - \frac{1}{\mathcal{Z}^2}\frac{\partial \mathcal{Z}_{\Lambda,K,h}}{\partial h_1}\frac{\partial \mathcal{Z}_{\Lambda,K,h}}{\partial h_2} \\
&= \langle s_1 s_2 \rangle - \langle s_1 \rangle \langle s_2 \rangle \\
&= \langle s_1 s_2 \rangle_c \text{ which is the connected correlation function.}
\end{aligned}
$$
(13)

Similarly, if there are four spins coupled to four independent magnetic field h_1, h_2, h_3, h_4 then

$$
\begin{aligned}
\left.\frac{\partial^4[-\beta F(K,\mathbf{h})]}{\partial h_1 \partial h_2 \partial h_3 \partial h_4}\right|_{h=0} &= \langle s_1 s_2 s_3 s_4 \rangle_c \\
&= \langle s_1 s_2 s_3 s_4 \rangle - \langle s_1 s_2 \rangle \langle s_3 s_4 \rangle - \langle s_1 s_3 \rangle \langle s_2 s_4 \rangle - \langle s_1 s_4 \rangle \langle s_2 s_3 \rangle.
\end{aligned}
$$
(14)

Since $h = 0$ in the end, there will be no correlators with an odd number of spins. Note that the previous derivations are valid even if s is replaced by some other variable, with eg. continuous values, coupled to the corresponding magnetic field in the Boltzmann weight, since the fact that $s = \pm 1$ was not used. The thermal averages may be computed in the case of non-zero zeros. In particular, if the field is uniform or zero. In that case, the h_i-derivatives are calculated and then set to $h_i = h$ or 0, for all i. Then, for example, if there were a uniform external field h, the connected correlation function is

$$(15) \qquad \langle s_1 s_2 \rangle - \langle s_1 \rangle \langle s_2 \rangle = \langle s_1 s_2 \rangle_c = \left.\frac{\partial^2[-\beta F(K,\mathbf{h})]}{\partial h_1 \partial h_2}\right|_{h_i = h,\,\forall i}.$$

The correlation function $\langle s_1 s_2 \rangle$ as a K-derivative,

$$
\begin{aligned}
\frac{\partial[-\beta F(K,\mathbf{h})]}{\partial K} &= \frac{\partial \log \mathcal{Z}_{\Lambda,K,h}}{\partial K} = \frac{1}{\mathcal{Z}_{\Lambda,K,h}}\frac{\partial \mathcal{Z}_{\Lambda,K,h}}{\partial K} = \frac{1}{\mathcal{Z}_{\Lambda,K,h}}\frac{\partial}{\partial K}\sum_{\mathbf{s}\in\Omega_\Lambda} e^{Ks_1 s_2 + h(s_1 + s_2)} \\
&= \frac{1}{\mathcal{Z}_{\Lambda,K,h}}\sum_{\mathcal{M}\times 2} s_1 s_2 e^{Ks_1 s_2 + h(s_1 + s_2)} = \langle s_1 s_2 \rangle \\
&= \frac{e^K \cosh 2h - e^{-K}}{e^K \cosh 2h + e^{-K}},
\end{aligned}
$$
(16)

which may interpreted as quantifying the average interaction energy, which happens to coincide with the correlation of neighbouring spins. In a model with more spins, correlations of not-neighbouring spins (eg. the

correlation between s_5 and s_{92}) cannot be obtained via a K-derivative. The distinction between the average inter-action energy and generic spin correlations is blurred in this toy-model.

Higher-order derivatives with respect to K or h (for uniform h) give additional information about fluctuactions about the mean. Thus,

$$(17) \qquad \chi = \frac{1}{2}\frac{\partial \langle M \rangle}{\partial h} = \frac{1}{4}\frac{\partial^2 [-\beta F]}{\partial h^2} = \langle M^2 \rangle_{\Lambda,K,h} - \langle M \rangle^2_{\Lambda,K,h}$$

measures the magnetization's fluctuation about its mean and is the magnetic susceptibility, and quantifies the rate of change of the average magnetization with the applied field.

One-dimensional Ising model. The standard way for solving the one-dimensional Ising model is to use the transfer matrix. The physical motivation behind it is very elegant: it consists on singling out a given spatial direction, say the x-direction. Then the partition function is built up by "evolving" the system from one value of $x = x_0$ to the next value $x = x_1$. Repeating this process yields the full partition function after all values of x are summed over. This method is not only useful for classical statistical mechanics, but by replacing the spatial evolution with a real time-evolution, yields a connection between quantum and classical statistical mechanics.

Now, any real system is finite, one may ask why we should focus on infinite systems. The answer lies that systems with a very large number of degrees of freedom look more like infinite systems than finite ones in the following sense. Phenomenologically, system at a temperature lower than some Curie temperature T_C will magnetize. But it can be shown rigorously that a finite system of Ising spins can never magnetize: if it is polarized one way (say the z-axis) for some time, it can jump to the opposite polarization after some time, so that on average it is not magnetized. This rigorous result can be reconciled with reality by noting that the time to flip can be as large as the age of the universe, so that in human time, magnetization is possible. The nice thing about infinite system is that remote possibilities are rendered impossible, making them a better model of real life.

Eventually, the thermodynamic limit $N \to \infty$ may be taken. The reason for lining up the lattice points on a vertical axis is that later it will become the time axis for a related problem. The energy of a given configuration may be written as

$$(18) \qquad E = -J \sum_{i=0}^{N-1} s_i s_{i+1}.$$

In the one-dimensional Ising model, there is only a single site at a fixed x-vale, so that the transfer matrix therefore "evolves" the system from one spin to the next. Therefore, consider two nearest-neighbour sites in the Ising model. There are, then, four configurations on these two sites, $++, +-, -+, --$. The Boltzmann weights $\mathbb{P}(\sigma_1 \sigma_2)$ for these four configurations are

$$(19) \qquad \mathbb{P}(++) = \mathbb{P}(--) = e^{\beta J} \text{ and } \mathbb{P}(+-) = \mathbb{P}(-+) = e^{-\beta J}.$$

Now consider for example an Ising model on three sites, with the simplest boundary conditions, where both spins σ_1 and σ_3 are both fixed at specific values. Summing over all values of the middle spin $\sigma_2 = \pm 1$, the partition function with both end spins fixed is

$$(20) \qquad \sum_{\sigma_2 = \pm 1} \mathbb{P}(+\sigma_2)\mathbb{P}(\sigma_2 +) = e^{2J} + e^{-2J}.$$

The general expression for the partition function of an ℓ-site Ising model with fixed boundary conditions can be found to be

If $N = 3$ then $\mathcal{Z}_3(\text{fixed}) = \sum_{\sigma_2 = \pm 1} \mathbb{P}(\sigma_1 \sigma_2) \mathbb{P}(\sigma_2 \sigma_3)$.

If $N = 4$ then $\mathcal{Z}_4(\text{fixed}) = \sum_{\substack{\sigma_2 = \pm 1 \\ \sigma_3 = \pm 1}} \mathbb{P}(\sigma_1 \sigma_2) \mathbb{P}(\sigma_2 \sigma_3) \mathbb{P}(\sigma_3 s \sigma_4)$.

These expression look exactly like matrix multiplications, written out in terms of the matrix elements. The matrix being multiplied is called the transfer matrix and includes all the nearest-neighbour interactions. For the one-dimensional Ising model it is given by

$$(21) \qquad \boldsymbol{\mathcal{T}} = \begin{pmatrix} \mathbb{P}(++) & \mathbb{P}(+-) \\ \mathbb{P}(-+) & \mathbb{P}(--) \end{pmatrix} = \begin{pmatrix} e^J & e^{-J} \\ e^{-J} & e^J \end{pmatrix}.$$

To compute the partition function with fixed boundaries conditions, from the transfer matrix, the boundary conditions are treated as vectors, on which the transfer $\boldsymbol{\mathcal{T}}$-matrix acts on. Here, the basis elements are the two spin values $+$ and $-$. These two vectors span a $\mathbb{C}^2$-vector space. The partition function of an N-site system with fixed boundary conditions can be found out to be

- if $N = 2$, $\mathcal{Z}_2(\text{fixed}) = \langle \sigma_1 | \boldsymbol{\mathcal{T}} | \sigma_2 \rangle = \mathbb{P}(\sigma_1 \sigma_2)$,
- if $N = 3$, $\mathcal{Z}_3(\text{fixed}) = \langle \sigma_1 | \boldsymbol{\mathcal{T}}^2 | \sigma_3 \rangle$,
- in general, $\mathcal{Z}_N(\text{fixed}) = \langle \sigma_1 | \boldsymbol{\mathcal{T}}^{N-1} | \sigma_N \rangle$.

Note that each time the transfer matrix is applied, it corresponds to adding one bond to the system. The transfer matrix now can easily be used to find analogous expressions for other boundary conditions. For example,

- for free boundary conditions, allowing the extreme-most spins to take on all allowed value, the partition function is given by

$$\mathcal{Z}_N(\text{free}) = \sum_{\substack{\sigma_1 = \pm 1 \\ \sigma_N = \pm 1}} \langle \sigma_1 | \boldsymbol{\mathcal{T}}^{N-1} | \sigma_N \rangle = \begin{pmatrix} 1 & 0 \end{pmatrix} \boldsymbol{\mathcal{T}}^{N-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \end{pmatrix} \boldsymbol{\mathcal{T}}^{N-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$(22) \qquad\qquad + \begin{pmatrix} 0 & 1 \end{pmatrix} \boldsymbol{\mathcal{T}}^{N-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} . \begin{pmatrix} 0 & 1 \end{pmatrix} \boldsymbol{\mathcal{T}}^{N-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 \end{pmatrix} \boldsymbol{\mathcal{T}}^{N-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

- for periodic boundary conditions, where an extra interaction $-J \sigma_N \sigma_1$ is added to the energy, including a bond between the two most-extreme spins. In this case, the model is translation invariant, shifting all the spins by one site mod N doesn't change the energy. The transfer $\boldsymbol{\mathcal{T}}$ must therefore evolve the first spin back to the N-th one. This is equivalent to having fixed boundary conditions on $N + 1$-sites where $\sigma_{N+1} = \sigma_1$, and then summing over both possible values, thus yielding

$$\mathcal{Z}_N(\text{periodic}) = \sum_{\sigma_1 = \pm 1} \langle \sigma_1 | \boldsymbol{\mathcal{T}} | \sigma_1 \rangle = \begin{pmatrix} 1 & 0 \end{pmatrix} \boldsymbol{\mathcal{T}}^N \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \end{pmatrix} \boldsymbol{\mathcal{T}}^N \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$(23) \qquad\qquad = \text{Tr}_{\mathbb{C}^2} \, \boldsymbol{\mathcal{T}}^N.$$

The partition functions with different boundary conditions are simply related, eg. $\mathcal{Z}_N(\text{periodic}) = \mathcal{Z}_N(++) + \mathcal{Z}_N(--)$. The usefulness of the transfer matrix in a one-dimensional system lies in that the partition function can be computed simply by diagonalizing the transfer matrix. For the one-dimensional Ising model, the diagonalization yields

$$(24) \qquad \boldsymbol{\mathcal{T}} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 2 \cosh(\beta J) & 0 \\ 0 & 2 \sinh(\beta J) \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}^{-1},$$

and given that the transfer matrix is hermitian, the eigenvalues are real. Then, the partition functions for the different boundary conditions can then be found

- For the free case, the $|\sigma_j\rangle$-vectors can be decomposed in terms of the $\mathcal{T}$- matrix's eigenvectors, which can be written as

$$v_\pm = \frac{1}{2}\Big(\begin{pmatrix} 1 & 1 \end{pmatrix}^{\mathrm{T}} + \begin{pmatrix} 1 & -1 \end{pmatrix}^{\mathrm{T}}\Big)$$

Therefore, plugging in these results in equation (22) yields,

$$(25) \qquad\qquad \mathcal{Z}_N(\text{free}) = \begin{pmatrix} 1 & 1 \end{pmatrix} \mathcal{T}^{N-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2(2\cosh\beta J)^{N-1}.$$

- For the periodic-boundary conditions, the partition function may be calculated as the sum of $\mathcal{Z}_N(++)$ and $\mathcal{Z}_N(--)$, which can be easily calculated from equation (24), as follows

$$(26) \qquad
\begin{aligned}
\mathcal{Z}_N(++) &= \begin{pmatrix} 1 & 0 \end{pmatrix} \mathcal{T}^N \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \tfrac{1}{2}\Big[\Big(2\cosh\beta J\Big)^N + \Big(2\sinh\beta J\Big)^N\Big] \\
\mathcal{Z}_N(+-) &= \begin{pmatrix} 1 & 0 \end{pmatrix} \mathcal{T}^N \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \tfrac{1}{2}\Big[\Big(2\cosh\beta J\Big)^N - \Big(2\sinh\beta J\Big)^N\Big] \\
\mathcal{Z}_N(-+) &= \begin{pmatrix} 0 & 1 \end{pmatrix} \mathcal{T}^N \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \tfrac{1}{2}\Big[\Big(2\cosh\beta J\Big)^N - \Big(2\sinh\beta J\Big)^N\Big] \\
\mathcal{Z}_N(--) &= \begin{pmatrix} 0 & 1 \end{pmatrix} \mathcal{T}^N \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \tfrac{1}{2}\Big[\Big(2\cosh\beta J\Big)^N + \Big(2\sinh\beta J\Big)^N\Big]
\end{aligned}$$

Then, given that $\mathcal{Z}_N(\text{periodic}) = \mathcal{Z}_N(++) + \mathcal{Z}_N(--)$, the periodic boundary condition partition functions is given by

$$\begin{aligned}
\mathcal{Z}_N(\text{periodic}) &= (2\cosh\beta J)^N + (2\sinh\beta J)^N \\
&= (2\cosh\beta J)^N\Big(1 + (2\tanh\beta J)^N\Big) \\
&\approx (2\cosh\beta J)^N \text{ for } N >> 1.
\end{aligned}$$

Note as well that the free energy per site is proportional to $\frac{\log \mathcal{Z}}{N}$, and is independent of boundary conditions when $N >> 1$. In effect, the free energy per site in the thermodynamic limit may be written as

$$(27) \qquad\qquad F(\beta, K) = -\lim_{N\to\infty} \frac{\log \mathcal{Z}}{N}.$$

It is straight forward to compute the two-point function in terms of these partition functions. Consider the following special cases.

- First, consider an Ising spin chain with free boundary conditions. Then, the two-point function between the two end-spins can be computed by summing over the four possible configurations. But these spins are opposite, then

$$(28) \qquad \langle \sigma_1 \sigma_N \rangle = \frac{\mathcal{Z}_N(++) + \mathcal{Z}_N(--) - \mathcal{Z}_N(+-) - \mathcal{Z}_N(-+)}{\mathcal{Z}_{\text{free}}} = (\tanh\beta J)^{N-1},$$

which falls off exponentially with distance, so even for small distances, the end spins are effectively uncorrelated.

- Now, consider an Ising chain with periodic boundary condtions. Then, the the two-point function between the j-th and the i-th spins, with $j > i$, may be written as

(29)
$$\langle \sigma_j \sigma_i \rangle_{\text{periodic}} = \frac{2\mathcal{Z}_{|j-i|}(++)\mathcal{Z}_{N-|j-i|}(--) + 2\mathcal{Z}_{|j-i|}(+-)\mathcal{Z}_{N-|j-i|}(-+)}{\mathcal{Z}_N \text{periodic}},$$

The general two-point correlator is similar and is given by

(30)
$$\langle \sigma_j \sigma_i \rangle_{j>i} = \frac{1}{\mathcal{Z}} \sum_k \sigma_j \sigma_i e^{\sum_k \beta J \sigma_k \sigma_{k+1}},$$

which measures how likely s_i and s_j are, on average, to point in the same direction. In any ferromagnetic system, the average will be positive for a pair of neighbouring spins since the Boltzmann weight is biased toward parallel values. Surely, if a spin can influence its neighbours to be parallel to it, they in turn will act similarly on their own neighbours, therefore long-range correlations may be present. In addition, if there is an external magnetic field present, it will enhance this correlation further. Naturally, if $j > i$ it holds that

$$\sigma_j \sigma_i = \sigma_i \sigma_j = \sigma_i \sigma_{i+1} \sigma_{i+1} \cdots \sigma_{j-1} \sigma_j = \mathsf{t}_i \mathsf{t}_{i+1} \cdots \mathsf{t}_{j-1} \text{ where } \mathsf{t}_i = \sigma_{i+1},$$

then

(31)
$$\langle \sigma_j \sigma_i \rangle = \langle \mathsf{t}_i \rangle \langle \mathsf{t}_{i+1} \rangle \cdots \langle \mathsf{t}_{j-1} \rangle.$$

Note that the answer factorizes over i since the Boltzmann weight factorizes over i as well, when written in terms of the t_i[4]. Then, the thermal average for any one t can be easily calculated as

(33)
$$\langle \mathsf{t} \rangle = \frac{1}{\mathcal{Z}} \sum_{\mathsf{t}=\pm 1} \mathsf{t} e^{E\{\mathsf{t}\}}$$
$$\langle \mathsf{t} \rangle = \frac{1 e^{\beta J} - 1 e^{-\cdot \beta J}}{e^{\beta J} + e^{-\cdot \beta J}} = \tanh \beta J \quad \Rightarrow \quad \langle \sigma_j \sigma_i \rangle = (\tanh \beta J)^{|j-i|} = \exp\left(|j-i| \log \tanh \beta J\right).$$

At any finite βJ, since $\tanh \beta J < 1$, $\langle \sigma_j \sigma_i \rangle \xrightarrow[|j-i| \to \infty]{} 0$. Only at $T = 0$ or $\beta J \to \infty$, when $\tanh \beta J = 1$, does the correlation not exponentially decay but remain flat. In other words, note that

(34)
$$\langle \sigma_j \sigma_i \rangle \sim e^{\frac{-|j-i|}{\zeta}} \text{ where } \zeta = -\frac{1}{\log \tanh \beta J} \text{ is the correlation length.}$$

Thus, the one-dimensional Ising model is in a disordered phase for any non-zero temperature. The zero-temperature case is pathological, since there is no meaning of equilibrium in a zero-temperature classical system.

The $\langle \sigma_j \sigma_i \rangle$-correlator depends on just the difference in coordinates, ie. it presents translational invariance. This is not a generic result for a finite chain, but a peculiarity of this model in particular. In general, for an $N+1$-site chain, correlations between two spins will generally depend on where the two points are in relation to the ends. On the other hand, for $N \to \infty$-models, translational invariance is expected to hold for correlations of spins far from the ends. In order for translational invariance to hold for a finite system, periodic boundary conditions must be imposed on the system. In said case, correlation functions will now only depend on the difference between the

---

[4]In effect, note that the partition function may be written as

(32)
$$\mathcal{Z} = \sum_{\sigma_i = \pm 1} \exp\left(\sum_{i=0}^{N-1} \beta J \sigma_i \sigma_{i+1}\right) = \sum_{\mathsf{t}_i = \pm 1} \exp\left(\sum_{i=0}^{N-1} \beta J \mathsf{t}_i\right)$$
$$= \sum_{\mathsf{t}_i = \pm 1} \prod_{i=0}^{N-1} e^{\beta J \mathsf{t}_i},$$

where the exponential has naturally factorized into a product over i.

two coordinates, but will not decay monotonically with separation since as one point start moving away from one extreme, it starts approaching the extreme spin from the other side.

The Monte-Carlo method. The partition function and two-point functions may be numerically calculated with the Monte-Carlo method. Consider the two point function,

$$(35) \qquad \langle \sigma_i \sigma_j \rangle_{\Lambda,\beta,h} = \frac{\sum_{\boldsymbol{\sigma} \in \Omega_\Lambda} \sigma_i \sigma_j e^{-\beta \mathbf{H}_{\Lambda,\beta,h}(\boldsymbol{\sigma})}}{\sum_{\boldsymbol{\sigma} \in \Omega_\Lambda} e^{-\beta \mathbf{H}_{\Lambda,\beta,h}(\boldsymbol{\sigma})}} = \sum_{\boldsymbol{\sigma} \in \Omega_\Lambda} \sigma_i \sigma_j(\boldsymbol{\sigma}) \mathbb{P}_{\Lambda,\beta,h}(\boldsymbol{\sigma}),$$

where $\sigma_i \sigma_j(\boldsymbol{\sigma})$ is the value of $\sigma_i \sigma_j$ at the $\boldsymbol{\sigma} \in \Omega_\Lambda$ configuration. The Monte-Carlo method is useful for numerically computing multidimensional integrals or sums. Each configuration will occur at a rate proportional to its Boltzmann weight. Then, for each configuration, the value of $\sigma_i \sigma_j$ can be computed. Since the configurations are already weighted, the arithmetic average of these numbers will yield the weighted thermal average. From this correlation function, the correlation length ε can be calculated. By the same method, any other quantity of interest may be calculated eg. the magnetization. This can be done for any problem with a real, positive, Boltzmann weight, not just for Ising models. One specific algorithm for calculating these weighted configurations is the **Metropolis Method**, as follows

Data: System's parameters: J, h, β
Result:
Start with some configuration $\boldsymbol{\sigma}_i \in \Omega_\Lambda$ with energy E_i; **while** *not at end of this document* **do**

　　Consider another configuration $\boldsymbol{\sigma}_j \in \Omega_\Lambda$ with energy E_j, obtained by changing some degrees of freedom;

　if $E_j < E_i$ **then**

　　｜ jump to j with unit probability;

　else

　　｜ jump to j with probability $e^{-\beta(E_j - E_i)}$

　end

end

Algorithm 1: Metropolis Algorithm

After a great number of iterations, the configurations appear with probabilities obeying

$$\frac{p(i)}{p(j)} = e^{-\beta(E_i - E_j)}.$$

In effect, the rate equation reads

$$\frac{dp(i)}{dt} = -p(i) \sum_j R(i \to j) + \sum_j p(j) R(j \to i),$$

where $R(i \to j)$ is the rate of jumping from the i-th configuration to the j-th configuration and vice-versa. In equilibrium or steady state, $\dot{p}(i) = 0$. One way to impose such constrain on the right-hand side is by considering a detailed balance approach, to ensure that the contribution from every value j is separately zero, rather than just the sum. It is necessary but not sufficient. In this case, this implies

$$(36) \qquad \frac{p(i)}{p(j)} = \frac{R(i \to j)}{R(i \to j)}.$$

2.2. **Quantum Spin Systems.** There is a deep correspondence between classical statistical mechanics and quantum statistical mechanics for spin-$\frac{1}{2}$ particles. In particular, the d-dimensional classical statistical mechanics problem may be mapped to a $d-1$-dimensional quantum statistical problem. The nature of the quantum variable will depend then on the nature of the classical variable. In general, the allowed values of the classical variable will correspond to the maximal set of simultanoes eigenvalues of the operators in the quantum problem (eg. σ_z).

The Hilbert space of a quantum spin is defined by choosing a representation for the spin operators. A representation of a Lie algebra is a set of three matrices satisfying the commutation relations of the $\mathfrak{su}(2)$-algebra. An **irreducible representation** is a set of matrices such that no unitary transformation $\mathcal{U}\mathbf{S}^a\mathcal{U}^\dagger$ block-diagonalizes all three matrices. It is known that for the $\mathfrak{su}(2)$-Lie algebra there is exactly one set (up to unitary transformations) of irreducible complex $n \times n$-matrices, for each integer n. It is customary to write $n = 2s + 1$ for all integers and half-integers s. A single spin-s quantum particle at a fixed point in space therefore has a Hilbert space $\mathbb{C}^{2s+1}$, so the matrices $\mathbf{S}^a$ are all $(2s + 1) \times (2s + 1)$. An orthonormal basis is given by the eigenstates of any one of the matrices.

- For $s = 0$, the matrices all consist of the number zero, thus this is the trivial representation.
- For $s = \frac{1}{2}$, the chosen basis is $\mathbf{S}^a = \frac{\hbar}{2}\sigma^a$, this is the fundamental representation.
- For $s = 1$, the matrices can be written to have entries $(\mathbf{S}^a)_{bc} = i\epsilon_{abc}$, yielding the adjoint representation.

In a given representation, an interesting invariant is given by the quadratic Casimir operator,

$$\mathbf{K} = \mathbf{S} \cdot \mathbf{S},$$

which commutes with each of the representation's generators. As a result, it must be proportional to the identity in a given irreducible representation. This is a fundamental consequence of Schur's lemma on theory of representations

Lemma 1. *Schur's lemma. Let $\mathbb{V}$ be a $\mathbb{C}$-vector space, associated with a finite-dimensional irreducible representation of an algebra $\mathfrak{A}$ over $\mathbb{C}$. Then, let $\phi : \mathbb{V} \to \mathbb{V}$ be a homorphism ie. $\phi(av) = a\phi(v), \forall a \in \mathfrak{A}, v \in \mathbb{V}$. Then, $\phi = \lambda \mathbb{1}_\mathbb{V}$.*

Spin chain types. Some of the most common examples of spin chains are treated as follows,

- The simplest example of a $SU(2)$-symmetric spin Hamiltonian is therefore the nearest-neighbor Heisenberg model, where

$$(37) \qquad \mathbf{H}_{\text{XXX}} = -J\sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j = -J\sum_{i=1}^{N} \left(\frac{1}{2}(\sigma_i^+\sigma_{i+1}^- + \sigma_i^-\sigma_{i+1}^+) + \frac{1}{4}\sigma_i^z\sigma_{i+1}^z \right),$$

where in the ferromagnetic case the diagonal terms in the Hamiltonian favour aligned spins, while in the antiferromagnetic case, the diagonal terms favour antialignement. The XXX-Hamiltonian commutes with the magnetization operator by construction. Note that Schur's lemma does not immediately apply since the magnetization operator is reducible.

- The XXZ model is a deformation of the Heisenberg model, breaking the $SU(2)$-model symmetry down to a $U(1)$-subgroup. The XXZ-hamiltonian reads

$$(38) \qquad \mathbf{H}_{\text{XXZ}} = -\sum_{\langle jk \rangle}(J_\perp(\mathbf{S}_j^x\mathbf{S}_k^x + \mathbf{S}_j^y\mathbf{S}_k^y) + J_z\mathbf{S}_j^z\mathbf{S}_k^z) = -\sum_{j=1}^{N} \left(J_\perp(\sigma_j^+\sigma_{j+1}^- + \sigma_j^-\sigma_{j+1}^+) + \frac{\Delta}{2}\sigma_j^z\sigma_{j+1}^z \right),$$

which is the most general $U(1)$-symmetric nearest-neighbor interaction for spin-$\frac{1}{2}$ particles.

Although this model is no longer $SU(2)$-symmetric, the magnetization operator commutes with the hamiltonian, thus generating a $U(1)$-symmetry. Even though the full $SU(2)$-symmetry is broken, the degrees of freedom are still typically referred to as spins. The sign of J_z alone determines whether the model is ferromagnetic, $J_z > 0$, or antiferromagnetic, $J_z < 0$, since the sign of $J_\perp$ is unimportant in any bipartite

lattice, redefining the states by changing the overall sign of the $\mathbf{S}^x$ and $\mathbf{S}^y$-operators on half the lattice sites, leaves the algebra unchanged but flips the sign of $J_\perp$. Therefore, the physically meaningful coupling is $\Delta = \frac{J_z}{|J_\perp|}$, so that $\Delta = \pm 1$ for the ferromagnetic and antiferromagnetic spin models, respectively. In the $\Delta \to \pm\infty$-limit, only the J_z remains, and the model is effectively classical. For the ferromagnetic case $J_z > 0$, all the spins simply line up with the maximum value of $\mathbf{M}^z$. In the antiferromagnetic case, $\Delta \to -\infty$ on a bipartite lattice, the spins take their maximum opposite values on every other site.

It is easy to check that this Hamiltonian commutes with the magnetization operator, and so preserves a $U(1) \times \mathbb{Z}_2$-symmetry. In a classical notion, the $U(1)$-symmetry corresponds to rotations around the z-axis, while the $\mathbb{Z}_2$ corresponds to flipping all the spins $\mathbf{S}_j^a \to -\mathbf{S}_j^a$.

- A still more general Hamiltonian is given by the XYZ-model, which only preserves the spin-flip symmetry.

None of these Hamiltonians correspond to a quantum version of the Ising model.

Ferromagnets and antiferromagnets. Unless geometric frustration is present, in classical systems, there is no significant difference between ferromagnets and antiferromagnets. Eg, with NN-interaction, geometric frustration occurs for lattices that are not bipartite. In a bipartite lattice, the sites can be divided into two sub-lattices such that nearest neighbours always belong to different sub-lattices. A nearest-neighbour antiferromagnetic interaction in a classical model on a bipartite lattice can typically be changed into a ferromagnetic one, by redefining the spin via a flip on all sites on one of the sub-lattices but not on the other (eg. $\uparrow \to \downarrow$ in the Ising model). The physics of such classical antiferromagnets is therefore essentially equivalent to that of the ferromagnets.

Antiferromagnetic quantum systems on non-bipartite lattices also exhibit interesting behaviour. But the interesting thing is that on bipartite lattices, there are a number of important differences between quantum ferromagnets and antiferromagnets.

Quantum Ferromagnets. Consider the Heisenberg interaction $-J\vec{\mathbf{S}}_1 \cdot \vec{\mathbf{S}}_2$ across a single bond. For spin-$\frac{1}{2}$ particles, this is a simple 4×4-matrix acting on the computational basis $\mathcal{B} = \{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\}$, yielding

$$\vec{\mathbf{S}}_1 \cdot \vec{\mathbf{S}}_2 = \frac{1}{4}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 4 & 0 \\ 0 & 4 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Diagonalizing this matrix will yield its eigenvalues and its eigenvectors. Given the Clebsh-Gordan decomposition rule $\frac{1}{2} \otimes \frac{1}{2} = 0 \oplus 1$, these eigenvectors can be grouped into the $s = 1$-triplet representation and the $s = 0$-singlet representation. In effect,

(39)
$$\text{triplet}: |\uparrow\uparrow\rangle, \frac{1}{\sqrt{2}}\Big(|\downarrow\uparrow\rangle + |\uparrow\downarrow\rangle\Big), |\downarrow\downarrow\rangle, \text{ with } \lambda_i = -\frac{J}{4}$$
$$\text{singlet}: \frac{1}{\sqrt{2}}\Big(|\downarrow\uparrow\rangle - |\uparrow\downarrow\rangle\Big), \text{ with } \mu_i = +\frac{J}{4}$$

It is simple to check that $\mathbf{K} = s(s+1)$ in both cases, and that while the singlet is annihilated by all three generators $\vec{\mathbf{S}}$, acting with $\mathbf{S}^+$ and $\mathbf{S}^-$ takes members of the triplet to each other. One important difference between the ferromagnetic ($J > 0$) and the antiferromagnetic ($J < 0$) models now arises,

- if $J > 0$, there are multiple ground states for the ferromagnet, each member of the triplet having minimum energy $-J$,
- but if $J < 0$, the antiferromagnetic ground state, the singlet, is unique. Moreover, the antiferromagnetic ground state is invariant under the $SU(2)$-symmetry, whereas the ferromagnetic ground states are not.

Additional insight comes from solving the Heisenberg model on a four-site chain, with its Hilbert space being 16-dimensional, by exploiting the model's inner symmetries. The magnetization operator $\mathbf{M}^z$ commutes with the Hamiltonian, which makes them simultaneously diagonalizable. For perioidic boundary conditions, the translational invariance plays a powerful role. The translation operator is defined by shifting the spins by one site modulo N, this is

$$
(40) \qquad \begin{aligned} \mathcal{T} &: \mathfrak{A}_i \to \mathfrak{A}_j \\ \mathcal{T}^{-1}\vec{\mathbf{S}}_i\mathcal{T} &= \vec{\mathbf{S}}_{i+1} \end{aligned} \quad \text{where } \mathfrak{A}_i, \mathfrak{A}_j \simeq \mathfrak{su}(2) \; \forall i, j.
$$

The Hamiltonian commutes with the translation operator when the boundary conditions are periodic, and commutes with the magnetization as well. Thus the Hamiltonian can be broken into blocks acting on states with a fixed eigenvalue of the traslation operator and of the magnetization.

Since $\mathcal{T}^N = \mathbb{1}$ for an N-site chain, the eigenvalues of $\mathcal{T}$ are $\lambda_i = e^{\frac{2\pi i n}{N}}, \forall n \in \mathbb{Z}$, from which the corresponding momentum can be defined as $k = \frac{2\pi n}{N}$. Then, the following holds

$$
\sum_{n \in \mathbb{Z}_{[0, N-1]}} e^{-\frac{2\pi i n}{N}} \mathcal{T}^n \left| s \right\rangle = \left| s \right\rangle + e^{-\frac{2\pi i}{N}} \left| s \right\rangle + \cdots + e^{-\frac{2\pi i (N-1)}{N}} \mathcal{T}^{N-1} \left| s \right\rangle.
$$

For example, for a four-site chain, the translational invariance diagonalizes the enire Hamiltonian save for $m = 0$ and $k = 0$ or $k = \pi$, since the corresponding eigenstates are

$$
(41) \qquad \begin{aligned} \left| A \right\rangle &= \frac{1}{\sqrt{2}}\left(\uparrow\downarrow\uparrow\downarrow + e^{-i\pi k}\downarrow\uparrow\downarrow\uparrow \right), \\ \left| B \right\rangle &= \frac{1}{2}\left(\left| \uparrow\uparrow\downarrow\downarrow \right\rangle + \left| \downarrow\uparrow\uparrow\downarrow \right\rangle + \cdots \right) \end{aligned} \quad \begin{aligned} &\text{with the hamiltonian} \\ &\text{on these} \\ &k = 0 \text{ and } k = \pi \\ &\text{states being} \end{aligned} \quad -J\begin{pmatrix} -1 & \sqrt{2}\cos\left(\frac{\pi k}{2}\right) \\ \sqrt{2}\cos\left(\frac{\pi k}{2}\right) & 0 \end{pmatrix}.
$$

These sector's eigenvalues are therefore $-J$ and $2J$ for $k = 0$ and $J = 0$ for $k = \pi$. Again organising the eigenstates into $\mathfrak{su}(2)$-multiplets yield the energy levels divided by $-J$, as follows

$$
(42) \qquad \begin{aligned} &\text{quintuplet}: 1 \\ &\text{triplets}: \cos(\pi k)(k \neq 0) \\ &\text{singlets}: -2, 0 \text{ for } k = 0, \pi \text{ respectively.} \end{aligned}
$$

This $\mathfrak{su}(2)$-invariance allows for the construction of the entire multiplet once one of the states is known.

As with the two-site chain, the ferromagnetic ground state is a multiplet, whereas the antiferromagnetic is a singlet. For a general N-site ferromagnet, the completely ferromagnetic states (all spins up or all spins down) are exact ground states of the Hamiltonian. This suggest using an order parameter for ferromagnetism. A ferromagnetic order parameter is simply the quantum analog of the magnetization, the expectation value of the z-component of the total spin is

$$
\left\langle \text{g.s.} \right| \mathbf{M}^{\mathbf{z}} \left| \text{g.s.} \right\rangle = \sum_{i=1}^{N} \text{g.s.} \left\langle \text{g.s.} \right| \mathbf{S}_{\mathbf{i}}^{\mathbf{z}} \left| \text{g.s.} \right\rangle,
$$

which is not a particularly great order parameter, since it vanishes if there is symmetry under spin flip, as with the classical case. A better example is $\left\langle \text{g.s.} \right| (\mathbf{M}^z)^2 \left| \text{g.s.} \right\rangle$.

The completely ferromagnetically ordered states, ie. all spins up or all spins down, are both $\mathbf{M}^z$-eigenvalues with maximum magnitude of eigenvalues, namely N^2, and are ground states of the Hamiltonian as well. The antiferromagnetic situtation is different. The staggered magnetization, or the **Néel order parameter**, is commonly used to understand antiferromagnetic order. It can only be defined on bipartite lattices. For example, consider the odd-sites on a sublattice, and the even-sites on the other, then the Néel operator is

$$\boldsymbol{\mathcal{N}} = \sum_{i=1}^{N} (-1)^i \mathbf{S}_i^z.$$

A Néel state has N^2-eigenvalues under $\boldsymbol{\mathcal{N}}^2$, ie. the spins are spin up on one sublattice and spin down on the other, like the state $|A\rangle$ which has four sites. Note, however, that the Néel operator does not commute with Hamiltonian. Then, a Néel state is not an eigenstate of the Heisenberg Hamiltonian. This is a huge difference between ferromagnets and antiferromagnets, even on bipartite lattices, not just on geometrically frustrated ones.

For example, for a four-site chain and $J < 0$, the ground state is a spin singlet and translationally invariant. The properties are thus quite typical of antiferromagnetic ground states. The diagonal terms are J and 0 for the translationally-invariant Néel state $|A\rangle$ and the other state, $|B\rangle$, respectively. This does indeed favour the Néel state. However, the off-diagonal terms in the Hamiltonian, mean

Hamiltonian in terms of Projectors. It is often useful to write a Hamiltonian as a sum of projection operators, especially if a form can be found where all the coefficients are positive.

.

9

2.3. **XXZ-model.** Consider the $d = 1$-XXZ-Heisenberg model, with its Hamiltonian (written in its longitudinal and transverse terms) as

$$\text{(43)} \qquad \mathbf{H}_{\text{XXZ}} = J \sum_{n=1}^{N} \left(\frac{1}{2}(\mathbf{S}_n^+ \mathbf{S}_{n+1}^- + \mathbf{S}_n^- \mathbf{S}_{n+1}^+) + \Delta \mathbf{S}_n^z \mathbf{S}_{n+1}^z \right),$$

where the effect of the transverse part, for spin-$\frac{1}{2}$ particles, is nothing but to interchange up and down spins, $|\uparrow\downarrow\rangle \Leftrightarrow |\downarrow\uparrow\rangle$, upto a factor of $\frac{1}{2}$. This model has a number of symmetries, leading to good quantum numbers such as wave vector $\mathbf{q}$ (translation), $\mathbf{M}^z$ (rotation about z-axis), $\mathbf{S}_{\text{tot}}$ (for general rotations, $|\Delta| = 1$) and parity (spin inversion). The most important link between theory and experiment are the spin correlation functions or response dynamical structure factors, which for a spin chain are defined as follows

$$\text{(44)} \qquad \begin{aligned} \boldsymbol{\mathcal{S}}^{\alpha,\alpha}(q,\omega) &= \sum_n \int dt\, e^{i(qn-\omega t)} \langle \mathbf{S}_n^\alpha \mathbf{S}_0^\alpha (t=0) \rangle, \\ \boldsymbol{\mathcal{S}}^{\alpha,\alpha}(q) &= \sum_n e^{iqn} \langle \mathbf{S}_n^\alpha \mathbf{S}_0^\alpha \rangle = \frac{1}{2\pi} \int d\omega\, \boldsymbol{\mathcal{S}}^{\alpha,\alpha}(q,\omega), \end{aligned}$$

where $\boldsymbol{\mathcal{S}}(q,\omega)$ determines the cross-section for scattering experiments as well as line shapes in NMR and ESR experiments. A useful sum rule is the total intensity, obtained by integrating $\boldsymbol{\mathcal{S}}(q)$ over frequency and wave vector,

$$\text{(45)} \qquad \frac{1}{4\pi^2} \int d\omega \boldsymbol{\mathcal{S}}^{\alpha,\alpha}(q,\omega) = \frac{1}{2\pi} \int dq \boldsymbol{\mathcal{S}}^{\alpha,\alpha}(q) = \langle (\mathbf{S}_0^\alpha)^2 \rangle,$$

which is simply equal to $\frac{1}{3}\boldsymbol{\mathcal{S}}(\boldsymbol{\mathcal{S}}+1)$ in the isotropic case.

The XXZ-Heisenberg chain is an important model to describe real material and is the most important paradigm of low-dimensional quantum magnetism, as it allows to introduce many scenarios, such as broken symmetry, the gapless Luttinger liquid, the Kosterlitz-Thouless phase transition, gapped and gapless excitation continua. Consider a Heisenberg XXZ chain supplemented by and external magnetic field.

Ferromagnetic phase: For $\Delta < -1$, the XXZ-chain is in the ferromagnetic Ising phase: the ground state is the saturated state, with all its spins aligned in either the z-direction or $-z$-direction. This entails that the classical ground state has magnetization $\mathbf{M}^z = \pm\frac{N}{2}$, where N is the number of lattice sites. This is a phase with broken symmetry since the ground state does not exhibit the discrete symmetry of spin reflection $\mathbf{S}^z \to -\mathbf{S}^z$, under which the Hamiltonian is invariant. Furthermore, when an external magnetic field in the z-axis is considered, and additional term is included into the Hamiltonian, namely a Zeeman interaction

$$\mathbf{H}_{\text{Zeeman}} = -g\mu_B H \sum_n \mathbf{S}^z.$$

Since the XXZ-Hamiltonian commutes with the magnetization operator, the external magnetic field results in an additional energy contribution $-g\mu_B H \mathbf{M}^z$, without affecting the wave-functions. The symmetry under spin reflection is lifted and the saturated ground state is stabilized. The low-lying excited states in the ferromagnetic phase are magnons with the total spin quantum number $\mathbf{M}_{\text{tot}}^z = \frac{1}{2}N - 1$ and the dispersion law, which is valid for general spin S

$$\text{(46)} \qquad \epsilon(q) = 2JS(1 - \cos(q) - (\Delta + 1)) + 2g\mu_B HS.$$

These states are exact eigenstates of the XXZ-Hamiltonian, In zero field, the excitation spectrum has a gap at $q = 0$ of magnitude $|\Delta| - 1$ for $\Delta < -1$. At $\Delta = -1$, the discrete spin-reflection symmetry generalizes to a continuous rotation symmetry and the spectrum thus becomes gapless. This is a consequence of Goldstone's theorem: the breaking of a continuous symmetry in the ground state results in the emergence of a gapless excitation mode. Whereas the ground state exhibit long range order, the large phase space available to the low-lying excitations in 1D leads to exponential decay of correlations at arbitrarily small finite temperatures following the Theorem of

Mermin and Wagner.

Two interesting consequences of equation (46) quickly arise, which can be generalized to any XXZ-type Hamiltonian conserving the total magnetization:

1) For sufficiently strong external magnetic field, the classical saturated state is forced to be the ground state for arbitrary value of Δ and the lowest excitations are exactly known. If the necessary magnetic fields are within experimentally accessible range, this can be used for an experimental determination of the exchange constants from the magnon dispersions (an example in 2D are recent neutron scattering experiments on Cs_2CuCl_4.

2) The ferromagnetic ground state becomes unstable when the lows spin wave frequency becomes negative. This allows to determine eg. the boundary of the ferromagnetic phase for $\Delta > -1$ in an external field as $H = H_c$ with $g\mu_B H_c = \Delta + 1$.

Néel phase: For $\Delta > +1$, the XXZ-chain is in the antiferromagnetic Ising or Néel phase with, in the thermodynamic limit, broken symmetry and one from 2 degenerate ground states, the $S = \frac{1}{2}$ remnants of the classical Néel states. The spatial period is $2a$, and the states are described in the reduced Brillouin zone with wave vectors $0 \le q \le \frac{\pi}{a}$. The ground states have zero magnetisation but finite sublattice magnetization,

$$\mathbf{N}^z = \sum_n (-1)^n \mathbf{S}_n^z,$$

and long range order in the corresponding correlation function. In contrast to the ferromagnet, however, quantum fluctuations prevent the order from being complete since the sublattice magnetization does not commute with the XXZ-Hamiltonian. For periodic boundary conditions and large but finite N, as in the situation in numerical approaches, the two ground states mix with energy separation $\propto e^{-\text{const}\times N}$, as $N \to \infty$. Then invariance under translation by the original lattice constant a is restored and the original Brillouin zone can be used, $\text{BZ} = \{q|0 \le q \le \frac{2\pi}{a}\}$.

The elementary excitations in the antiferromagnetic Ising phase are described most clearly close to the Ising limit $\delta \to \infty$, starting from one of the two ideal Néel states: turning around one spin breaks two bonds and leads to a state with energy Δ, degenerate with all state resulting from turning around an arbitrary number of subsequent spins. These states have total magnetization $\mathbf{M}^z = \pm 1$, response 0 for an odd, and response even number of turned spins. They are appropriately called two-domain wall states since each of the two broken bonds mediates between two different Néel states. They are appropriately called two-domain wall states, since each of the two broken bonds mediates between two different Néel states. The total number of these states is $N(N-1)$,

- there are $\frac{N^2}{4}$ states with total magnetization $\mathbf{M}^z = 1$ and $\mathbf{M}^z = -1$ (number of turned spins odd),
- and $\frac{N^2}{2} - N$ states with null magnetization (number of spins turned even).

These states are no more eigenstates when Δ^{-1} is finite, but for $\Delta^{-1} << 1$, they can be dealt with in perturbation theory, leading to the excitation spectrum in the first order in $\frac{1}{\Delta}$

$$\omega(q, k) = \Delta + 2\cos q \cos 2\Phi = \epsilon\left(\frac{\epsilon}{2} + \Phi\right) + \epsilon\left(\frac{\epsilon}{2} - \Phi\right),$$

with $\epsilon(k) = \frac{\Delta}{2} + \cos 2k$, where q is the total momentum and takes values $q = \frac{2\pi\ell}{N}, \ell \in \mathbb{Z}_{[1,\frac{N}{2}]}$, Φ is the wave vector related to the superposition of domain walls with different distances and for $\mathbf{M}^z = \pm 1$ takes values $\Phi = \frac{m\pi}{N+2}$ with $m \in \mathbb{Z}_{[1,\frac{N}{2}]}$. Φ is, in essence, a relative momentum.

XY phase. For $-1 < \Delta < 1$ and zero external field, the XXZ-chain undergoes an XY phase, characterized by uniaxial symmetry of the easy-plane type and a gapless excitation continuum. Whereas the full analysis of this phase for general Δ, requires the use of bosonization and Bethe ansatz, a somewhat simpler approach is based on the mapping of $S = \frac{1}{2}$ spin operators in 1D to spinless fermions via the non-local Jordan Wigner transformations, wherein the mapping reads:

$$(47) \qquad \mathbf{S}_n^+ = f_n^\dagger e^{i\pi \sum_{j=1} f_p^\dagger f_p}, \qquad\qquad \mathbf{S}_n^z = f_n^\dagger f_n - \frac{1}{2}.$$

When a fermion is at site n, the spin projection is thus $\mathbf{S}_n^z = \pm\frac{1}{2}$. Thus, in terms of the fermionic operators, the XXZ-Hamiltonian reads

$$(48) \qquad \mathbf{H}_{\text{XXZ}} = J \sum_n \left[\frac{1}{2}(f_n^\dagger f_{n+1} + f_{n+1}^\dagger f_n) + \Delta \left(f_n^\dagger f_n - \frac{1}{2} \right) \left(f_{n+1}^\dagger f_{n+1} - \frac{1}{2} \right) - g\mu_B H \left(f_n^\dagger f_n - \frac{1}{2} \right) \right].$$

For general Δ, the XXZ-chain is thus equivalent to an interacting 1D fermion system. In the $\Delta = 0$ case, this model simplifies itself to an integrable XX-chain, wherein the fermion chain becomes noninteracting and has a simpler analytical solution.

XX-model: $\Delta = 0$. Consider the XX-Heisenberg model, with its Hamiltonian given in terms of the traditional $\frac{1}{2}$-spin operators ie.

$$(49) \qquad \mathbf{H}_{\text{XX}} = J \sum_{i=1}^L (\mathbf{S}_j^x \mathbf{S}_{j+1}^x + \mathbf{S}_j^y \mathbf{S}_{j+1}^y) - \lambda \sum_{j=1}^L \mathbf{S}_j^z,$$

which describes interacting spins in a one-dimensional chain, with periodic boundary conditions. (49)'s first terms represents nearest neighbour interactions in the x and $y-$directions interactions, with J being either positive or negative and quantifying the strength and type of interactions, while the second term represents a magnetic field of strength λ, applied in the z-direction of the spins.

Exact solution

In order to solve this problem, it is necessary to rewrite (49) and apply a Jordan-Wigner transformation, mapping the spin problem into a fermionic problem. But first, it is convenient to write the spin-operators in terms of the raising and lowering $\mathfrak{su}(2)$-operators, ie.

$$\mathbf{S}_j^\pm = \mathbf{S}_j^x \pm i\mathbf{S}_j^y \Rightarrow \begin{array}{c} \mathbf{S}_j^x = \frac{1}{2}(\mathbf{S}_j^+ + \mathbf{S}_j^-), \\[4pt] \mathbf{S}_j^y = \frac{1}{2i}(\mathbf{S}_j^+ - \mathbf{S}_j^-). \end{array} \Rightarrow \mathbf{H}_{\text{XXZ}} = \frac{J}{2} \sum_{i=1}^L (\mathbf{S}_j^+ \mathbf{S}_{j+1}^- + \mathbf{S}_j^- \mathbf{S}_{j+1}^+) - \lambda \sum_{j=1}^L \mathbf{S}_j^z$$

where the interacting terms in the first summation, flip neighboring spins if said spins are anti-aligned[5]. In addition, the XX-Hamiltonian has a total magnetization symmetry, since the Hamiltonian given by (2.3) conmutes with the magnetization operator.

---

[5]In effect, consider for example, a two-spin problem. Then, the interaction term is given by

$$\mathbf{S}_1^+ \mathbf{S}_2^- + \mathbf{S}_1^- \mathbf{S}_2^+,$$

and consider an anti-aligned state $|\downarrow\uparrow\rangle$. Then, the action of the previous two-spin operator over this state yields

$$(\mathbf{S}_1^+ \mathbf{S}_2^- + \mathbf{S}_1^- \mathbf{S}_2^+) |\downarrow\uparrow\rangle = |\uparrow\downarrow\rangle + 0,$$

since $\mathbf{S}_1^- \mathbf{S}_2^+$ destroys the state. Similarly, $(\mathbf{S}_1^+ \mathbf{S}_2^- + \mathbf{S}_1^- \mathbf{S}_2^+) |\uparrow\downarrow\rangle = |\downarrow\uparrow\rangle$. However, note that, should both spins be either up or down, the state remain invariant under the action of the two-spin operator.

$$(50) \qquad (\mathbf{S}_1^+ \mathbf{S}_2^- + \mathbf{S}_1^- \mathbf{S}_2^+) |\downarrow\downarrow\rangle = |\downarrow\downarrow\rangle \text{ and } (\mathbf{S}_1^+ \mathbf{S}_2^- + \mathbf{S}_1^- \mathbf{S}_2^+) |\uparrow\uparrow\rangle = |\uparrow\uparrow\rangle$$

Under the Jordan-Wigner map, nearest-neighbours spin flipping is translated into to nearest-neighbours fermionic hopping, ie. $\mathbf{S}_j^+\mathbf{S}_{j+1}^- = f_j^\dagger f_{j+1}$ and $\mathbf{S}_j^-\mathbf{S}_{j+1}^+ = f_{j+1}^\dagger f_j$. However, due to the boundary conditions' periodicity, the XX-Hamiltonian cannot be rewritten as a fermionic model yet since (it will contain an additional boundary term), for example, the fermionic counterparts to the $\mathbf{S}_L^+\mathbf{S}_1^-$ interaction are highly non-local operators and are not desirable. Indeed, under the Jordan-Wigner mapping

$$\mathbf{S}_L^+\mathbf{S}_1^- = f_L^\dagger \exp\left(-i\pi\sum_{\ell=1}^{L-1} f_\ell^\dagger f_\ell\right) f_1 \Rightarrow \begin{array}{l} \mathbf{S}_L^+\mathbf{S}_1^- = \mathcal{Q} f_L^\dagger f_1 \\ \mathbf{S}_L^-\mathbf{S}_1^+ = \mathcal{Q} f_1^\dagger f_L, \end{array}$$

then (2.3) can be rewritten as

$$(51) \qquad \mathbf{H}_{\mathbb{XX}} = \frac{J}{2}\sum_{i=1}^{L-1}\left(f_j^\dagger f_{j+1} + f_{j+1}^\dagger f_j\right) - \lambda\sum_{j=1}^{L}\left(f_j^\dagger f_j - \frac{1}{2}\right) + \frac{J}{2}\mathcal{Q}(f_L^\dagger f_1 + f_1^\dagger f_L),$$

where the first term accounts for fermionic nearest-neighbour hopping, the second term accounts for the magnetic field, and the third term being the non-local boundary term. Note that this fermionic Hamiltonian hasn't got any type of boundary conditions, since the L-lattice site is disconnected in any way whatsoever from the first lattice site. Then, the standard procedure is to add and subtract terms from the Hamiltonian, so that the nearest-neighbour hopping term in (51) can also have periodic boundary conditions, thus yielding

$$(52) \qquad \mathbf{H}_{\mathbb{XX}} = \frac{J}{2}\sum_{i=1}^{L}\left(f_j^\dagger f_{j+1} + f_{j+1}^\dagger f_j\right) - \lambda\sum_{j=1}^{L}\left(f_j^\dagger f_j - \frac{1}{2}\right) + \frac{J}{2}(\mathcal{Q}-1)(f_L^\dagger f_1 + f_1^\dagger f_L),$$

where now the fermionic hopping term has the standard boundary conditions. The third term, since it does not involve any type of summation over lattice sites, only contributes at $\mathcal{O}\left(\frac{1}{L}\right)$-order to any microscopic quantity. In the thermodynamic limit, this non-local term can be dropped, thus yielding an $\mathcal{O}(L)$-Hamiltonian given by

$$(53) \qquad \mathbf{H}_{\mathbb{XX}} = \frac{J}{2}\sum_{i=1}^{L}\left(f_j^\dagger f_{j+1} + f_{j+1}^\dagger f_j - \lambda f_j^\dagger f_j\right) + \frac{\lambda L}{2},$$

which is now fully cyclic and where its operators obey fermionic algebras. This Hamiltonian can then be diagonalized via a discrete Fourier transform on the fermionic operators

$$(54) \qquad f_j = \frac{1}{\sqrt{L}}\sum_{\substack{k=2\pi m/L \\ m\in\mathbb{Z}_{[1,L]}}} e^{ijk}d_k, \qquad\qquad f_j^\dagger = \frac{1}{\sqrt{L}}\sum_{\substack{k=2\pi m/L \\ m\in\mathbb{Z}_{[1,L]}}} e^{-ijk}d_k^\dagger,$$

with the inverse transformation given by

$$(55) \qquad d_k = \frac{1}{\sqrt{L}}\sum_{j=1}^{L} e^{-ikj}f_j \qquad\qquad d_k^\dagger = \frac{1}{\sqrt{L}}\sum_{j=1}^{L} e^{ikj}f_j^\dagger.$$

Note that the d_k-operators follow the standard fermionic anticonmutation algebra. Note as well, that the f_j-vacuum state, defined such that $f_j|0\rangle_f = 0$, $\forall j$, is the same as the d_k-vacuum state, defined such that $d_j|0\rangle_d = 0$, $\forall k$, ie. $|0\rangle_f = |0\rangle_d$. Another important relationship is the Fourier transform's consistency condition, ie.

$$\sum_{j=1}^{L} e^{i(k-q)j} = L\delta_{kq}.$$

Under the Fourier transform, (53)'s terms are mapped as follows

$$\sum_{j=1}^{L} f_j^\dagger f_{j+1} = \sum_{j=1}^{L} \frac{1}{L} \sum_{k,q} e^{-ikj} e^{iq(j+1)} d_k^\dagger d_q = \sum_{j=1}^{L} \frac{1}{L} \sum_{k,q} e^{i(q-k)j} e^{iq} d_k^\dagger d_q$$

$$= \sum_{k,q} \frac{1}{L} e^{iq} \delta_{qk} d_k^\dagger d_q = \sum_{k} e^{ik} d_k^\dagger d_k$$

$$\sum_{j=1}^{L} f_{j+1}^\dagger f_j = \sum_{j=1}^{L} \frac{1}{L} \sum_{k,q} e^{-ik(j+1)} e^{iqj} d_k^\dagger d_q = \sum_{j=1}^{L} \frac{1}{L} \sum_{k,q} e^{i(q-k)j} e^{-ik} d_k^\dagger d_q$$

$$= \sum_{k,q} \frac{1}{L} e^{-ik} \delta_{qk} d_k^\dagger d_q = \sum_{k} e^{-ik} d_k^\dagger d_k$$

$$\sum_{j=1}^{L} f_j^\dagger f_j = \sum_{j=1}^{L} \frac{1}{L} \sum_{k,q} e^{-ikj} e^{iqj} d_k^\dagger d_q = \sum_{j=1}^{L} \frac{1}{L} \sum_{k,q} e^{i(q-k)j} d_k^\dagger d_q$$

$$= \sum_{k,q} \frac{1}{L} \delta_{qk} d_k^\dagger d_q = \sum_{k} d_k^\dagger d_k$$

$$\Rightarrow \mathbf{H}_{\mathbb{XX}} = \frac{J}{2} \sum_{k} \left(e^{ik} d_k^\dagger d_k + e^{-ik} d_k^\dagger d_k - \lambda d_k^\dagger d_k \right) + \frac{\lambda L}{2} = \sum_{k} \left(J \cos k - \lambda \right) d_k^\dagger d_k + \frac{\lambda L}{2},$$

which can be rewritten as

(56)

$$\mathbf{H}_{\mathbb{XX}} = \sum_{k} \epsilon_k d_k^\dagger d_k + \frac{\lambda L}{2}$$

with the eigenvalues being given by $\epsilon_k = J \cos k - \lambda + \frac{\lambda L}{2}$ and the eigenvectors being given by $|E_n\rangle = \prod_n (d_k^\dagger)^n$, with eigenvalue $E_n = \sum_n \epsilon_n$

where $k = \frac{2\pi m}{L}$, $m = -\frac{L}{2}+1, \cdots \frac{L}{2}$, ie. $k \in (-\pi, \pi]$, this is called the first Brillouin zone. Thus the problem has been solved. As for its thermal properties, since this is a fermionic model, the fermions will obey the Fermi-Dirac distribution, ie.

(57)
$$\mathcal{N}_{jk} = \langle d_j^\dagger d_k \rangle_{\text{th}} = \frac{1}{1 + e^{\beta \epsilon_k + \mu}} \delta_{jk}$$

Quantum Phase Transition

Note that each of the fermion states can be either occupied or vacant, corresponding to the dimension 2^N of the Hilbert space for N spins with $S = \frac{1}{2}$. The ground state, the state with the lowest energy, has all levels with $\epsilon(k) \leq 0$ and can then be defined such that

(58)
$$\langle E_G| \, d_k^\dagger d_k \, |E_G\rangle = 1 \text{ if } \epsilon_k \leq 0 \text{ and } \langle E_G| \, d_k^\dagger d_k \, |E_G\rangle = 0 \text{ if } \epsilon_k > 0.$$

Then the groundstate energy is given by

(59)
$$E_G = \sum_k \epsilon_k \, \langle E_G| \, d_k^\dagger d_k \, |E_G\rangle .$$

Now, there exists a momenta k_F such that $\epsilon_{k_F} = 0$, with k_F being the Fermi momentum, ie. $\epsilon_{k_f} = J \cos k_f - \lambda$. The previous equation depends both on the coupling strength J and the magnetic field strength λ. From the previous equation some special cases arise, namely

- if $\lambda > J$, all fermions levels are occupied, which lead to maximum positive magnetization. This entails that $k_F = 0$.
- If $\lambda < -J$, all fermion levels are vacant, leading to maximum negative magnetization. This implies a $k_F = \pi$ Fermi momentum.
- If $|\lambda| \leq J$, the Fermi momentum is $k_F = \arccos\left(\frac{\lambda}{J}\right)$.

Therefore, the system groundstate will be defined by the following relations,

(60)
$$\langle E_G| \, d_k^\dagger d_k \, |E_G\rangle = 1 \qquad \text{and} \qquad \langle E_G| \, d_k^\dagger d_k \, |E_G\rangle = 0$$
$$\text{if } k \in [-\pi, -k_F] \cup [k_F, \pi], \qquad \text{if } k \in [-k_F, k_F].$$

Note that the periodic boundary conditions in spin space are modified by the Jordan-Wigner transformation, the boundary term in the Hamiltonian explicitly depends on the fermion number L and leads to different Hamiltonians for the two subspaces of even or odd fermion number. For fixed fermion number, this reduces to different sets of allowed fermion momenta k, if the total number of spins L is even, the allowed values of fermion momenta are given by $k_n = \frac{2\pi m}{L}$, where the m-numbers are integer/half-odd integers if the number of fermions $L_f = \mathbf{M}^z + \frac{N}{2}$ is odd/even. The total momentum for the ground states is thus $p_F = L_f \pi$[6].

The total magnetization operator is defined as the sum of the $\mathbf{S}_z$ operators, along the entire chain ie. $\mathbf{M}^z = \sum_{j=1} \mathbf{S}_j^z$. Under the Jordan-Wigner transformation, $\mathbf{S}_j^z = f_j^\dagger f_j - \frac{1}{2}$. In turn, these fermionic number operators can be mapped to fermionic number operators acting on momentum space, ie. $\sum_{j=1}^L f_j^\dagger f_j = \sum_k d_k^\dagger d_k$. This entails

(61)
$$\mathbf{M}^z = \sum_k \left(d_k^\dagger d_k - \frac{1}{2}\right).$$

Then the total magnetization operator's expectation value for the groundstate can be calculated. Given the system's fermionic nature, momentum eigenmodes are either occupied or unoccupied, contributing only one or zeros respectively. Then, this expectation value yields

(62)
$$\langle E_G| \, \mathbf{M}^z \, |E_G\rangle = \sum_{k=-\pi}^{-k_F} 1 + \sum_{k=k_F}^{\pi} 1 - \frac{1}{2}\sum_{k=-\pi}^{\pi} 1 = \frac{1}{2}\left(\sum_{k=-\pi}^{-k_F} 1 + \sum_{k=k_F}^{\pi} 1 - \sum_{k=-k_F}^{k_F} 1 \right),$$

where the total magnetization expectation value is just half the difference of two competing summations, related to the negative and positive eigenmodes. Defining the total magnetization per site operator, $\mathbf{m}^z = \frac{1}{L}\mathbf{M}^z$. Now, as

---

[6]The same two sets of k-values are found via the Bethe ansatz solution of the XXZ-chain. The complication of two different Hilbert spaces is avoided entirely with free boundar conditions, giving up translational symmetry.

the chain gets progressively bigger, with more and more lattice sites, the summations in (62) can be approximated by Riemannian integrals, ie.

$$\sum_k f(k) = \frac{1}{\Delta k} \sum_k f(k) \Delta k \stackrel{L \to \infty}{\Rightarrow} \frac{L}{2\pi} \int_{k \in \mathbb{BZ}} dk \, f(k) \text{ where } \Delta k = \frac{2\pi}{L}.$$

Using this trick, (62) can be rewritten as

$$(63) \qquad \langle \mathbf{m}^z \rangle_G = \frac{1}{4\pi} \left(\int_{-\pi}^{-k_F} dk + \int_{k_F}^{\pi} dk - \int_{-k_F}^{k_F} dk \right) = \frac{1}{4\pi} \left(-k_F + \pi + \pi - k_F - k_F + k_F \right)$$

$$(64) \qquad = \frac{1}{2} - \frac{k_F}{\pi},$$

which yields the total magnetization per lattice site. Note that at $\lambda < J$, $k_F = \pi$ whereby there is total polarization ie. $\langle \mathbf{m}^z \rangle_G = -\frac{1}{2}$. Similarly, for $\lambda > J$, $k_F = 0$ and then $\langle \mathbf{m}^z \rangle_G = \frac{1}{2}$. However, if $|\lambda| \leq J$, then $k_F = \arccos\left(\frac{\lambda}{J}\right)$, which then yields a total magnetization per lattice site given by

$$(65) \qquad \langle \mathbf{m}^z \rangle_G = \begin{cases} \frac{1}{2} - \frac{1}{\pi} \arccos\left(\frac{\lambda}{J}\right) & \text{if } |\lambda| \leq J \\[2mm] -\frac{1}{2} & \text{if } \lambda < -J \\[2mm] \frac{1}{2} & \text{if } \lambda > J \end{cases}$$

Note that for $\lambda < J$, maximum negative magnetization is obtained, while for $\lambda > J$ maximum positive magnetization is obtained. Another interesting quantity is the total magnetic susceptibility, which can be written in terms of the total magnetization per lattice site as follows

$$(66) \qquad \chi = \left(\frac{\partial \langle \mathbf{m}^z \rangle_G}{\partial \lambda} \right) \Rightarrow \chi = \begin{cases} \frac{1}{J\pi\sqrt{1 - \frac{\lambda^2}{J^2}}} & \text{if } |\lambda| \leq J \\[2mm] 0 & \text{if } \lambda < -J \\[2mm] 0 & \text{if } \lambda > J \end{cases},$$

which is divergent if $\lambda = J$. This shows that the XX-Heisenberg model has a second order (since there is no latent heat involved it can't be a first order phase transition) quantum phase transition, with an associated power law with a critical exponent.

Static correlation functions for the XX-model can be calculated for the discrete system, without taking the continuum limit. The longitudinal correlation function in the ground state is given by

$$\langle 0 | \mathbf{S}_n^z \mathbf{S}_0^z | 0 \rangle = \begin{cases} -\frac{1}{4} \left(\frac{2}{\pi n} \right)^2 & \text{for } n \in 2\mathbb{N} + 1 \\ 0 & \text{for } n \in 2\mathbb{N} \end{cases}$$

The transverse correlation function is expressed as a product of two $\frac{n}{2} \times \frac{n}{2}$ determinants. An explicit expression is available only for the asymptotic behaviour

$$\langle 0 | \mathbf{S}_n^x \mathbf{S}_0^x | 0 \rangle = \langle 0 | \mathbf{S}_n^y \mathbf{S}_0^y | 0 \rangle \sim \frac{C}{\sqrt{n}}, \quad C \approx 0.5884....$$

Dynamic correlation functions cannot be obtained at the same level of rigour as static ones since they involve transitions between states in different Hilbert spaces (one for even and another for odd fermion number).

Quantities of experimental relevance can be easily calculated from the exact expression for the free energy in terms of the basic fermion dispersion,

$$(67) \qquad F = -Lk_BT\left[\log 2 + \frac{2}{\pi} \int_{\mathbb{R}_{[0,\frac{\pi}{2}]}} dk \log \cosh \left(\frac{\epsilon(k)}{2k_BT} \right) \right].$$

An important quantity is the specific heat whose low-temperature behaviour is linear in T:

$$C(T) \approx \frac{\pi T}{6v_F}, \text{ where } v_F = \frac{\partial \epsilon(k)}{\partial k}\bigg|_{k=k_F} = J \text{ is the Fermi velocity.}$$

Low-lying excitations are also simply described in the fermion picture: they are either obtained by adding or removing fermions, thus changing the total magnetization by one unity and adding or removing the energy $\epsilon(k)$, or particle-hole excitations which do not change the total magnetization. Creating a particle-hole excitation involves moving a fermion with momentum k_i inside the Fermi sea t o some momentum k_f, outside of the Fermi sea. It is clear that moving a fermion just across the Fermi point costs arbitrarily low energy since the excitation spectrum is gapless. It is easily seen that for a given total momentum $q = k_f - k_i$, a finite range of excitation energies is possible, thus the spectrum of particle-hole excitations is a continuum with the initial momentum $k = ki$ as an internal degree of freedom,

$$\omega(q, k) = \epsilon(k + q) - \epsilon(k).$$

For $\Delta \neq 0$ the interacting fermion Hamiltonian can be treated in perturbation theory. From this approach, and more generally from the Bethe ansatz and field-theoretic methods, it is established that the behaviour for $-1 < \Delta < 1$ is qualitatively the same as the free fermion limit $\Delta = 0$: the excitation spectrum is gapless, a Fermi point exists and correlation functions show a power-law behaviour. The Heisenberg chain is in the XY regime, and thus is in a critical phase. The phase is equivalent to the so-called Tomonaga-Luttinger liquid. The fermion dispersion, upto first order in Δ, is obtained by direct perturbation theory starting from the free fermion limit (in unit of J),

$$\epsilon(k) = \Delta - \lambda + \cos q - \frac{2\Delta}{\pi}\Theta(1 - \lambda)\left[\arccos \lambda - \frac{\cos q}{(1 - \lambda^2)^{\frac{1}{2}}} \right].$$

For the more general XYZ-chain, wherein the Hamiltonian reads

$$(68) \qquad \mathbf{H}_{\mathrm{XYZ}} = J \sum_{n=1}^{N} \left((1+\gamma)\mathbf{S}_n^x\mathbf{S}_{n+1}^x + (1-\gamma)\mathbf{S}_n^y\mathbf{S}_{n+1}^y + \Delta\mathbf{S}_n^z\mathbf{S}_{n+1}^z \right) - \lambda \sum_{n=1}^{N} \mathbf{S}_n,$$

the rotational symmetry in the xy-plane is broken and a unique preferred direction in spin space exists: $\Delta = 0$ continues to result in a free fermion system, but the basic fermion dispersion acquires a gap and the ground state correlation functions $\langle 0|\, \mathbf{S}_n^x\mathbf{S}_0^x\, |0\rangle$ develops long range order.

$\Delta \approx 1$: **The isotropic Heisenberg Antiferromagnet**. The most interesting regime for the XXZ chain is $\Delta \approx 1$, ie. the vicinity of the Isotropic Heisenberg antiferromagnet (HAF), wherein the ground state energy is given by

$$E_0 = -NJ\log 2.$$

2.4. Bethe ansatz.

Heisenberg XXX-chain. Consider a one-dimensional lattice with N lattice sites and a $\frac{1}{2}$-spin particle positioned at every lattice site, which have a nearest neighbour spin-spin interaction. Each particles either has spin up or spin down, generating a two-dimensional local Hilbert space $\mathbb{V}_n$. Strictly speaking there are two possible topologies for a one-dimensional chain, open or closed. Consider the XXX-Hamiltonian with closed topology, given by

$$
\begin{aligned}
\mathbf{H} &= J \sum_{n=1}^{N} \left(\vec{\mathbf{S}}_n \cdot \vec{\mathbf{S}}_{n+1} - \frac{1}{4} \mathbb{1}^{\otimes N} \right) \\
&= \frac{J}{4} \sum_{n=1}^{N} \left(\vec{\sigma}_n \cdot \vec{\sigma}_{n+1} - \frac{1}{4} \mathbb{1}^{\otimes N} \right).
\end{aligned}
$$

(69)

Note that $J < 0$ models a ferromagnetic state while $J > 0$ models an anti-ferromagnetic state. A ferromagnetic state refers to a state where all the spins are alligned while an anti-ferromagnetic state refers to a state where adjacent spins are anti-aligned within the domain.

The main point of the Bethe solution is that XXX Heisenberg Hamiltonian, which defines the model, can be expressed in terms of a complex-valued $\mathcal{R}$-matrix which is a solution of the Yang-Baxter equation, thus leading to the model's integrability. The most convenient approach is to construct a transfer $\mathcal{T}$-matrix -a one parameter commutative family of operators acting on the full state space of the Heisenberg spin chain.

Some fundamentals of Matrix Theory. Let $\mathbb{V}$ be an n-dimensional $\mathbb{C}$-vector space and let $k \in \mathbb{1}_n$ such that $\in \mathbb{N}_{[0,k]}$. Let $\mathfrak{J} = \mathbb{1}_n / \mathbb{1}_k = \{k+1, \cdots, n\}$. Let $A \in \mathrm{GL}(n, \mathbb{C})$, which can be written in block notation as

(70)
$$
A = \begin{pmatrix} A_{\mathcal{I}} & A_{\mathcal{I}\mathfrak{J}} \\ A_{\mathfrak{J}\mathcal{I}} & A_{\mathfrak{J}} \end{pmatrix} \begin{matrix} \mathbb{C}^k \\ \mathbb{C}^{n-k} \end{matrix} \quad \text{where} \quad \begin{matrix} A_{\mathcal{I}} \in \mathrm{GL}(k, \mathbb{C}), & A_{\mathcal{I}\mathfrak{J}} \in \mathrm{GL}(k \to n-k, \mathbb{C}), \\ A_{\mathfrak{J}\mathcal{I}} \in \mathrm{GL}(n-k \to k, \mathbb{C}), & A_{\mathfrak{J}} \in \mathrm{GL}(n-k, \mathbb{C}), \end{matrix}
$$

This notation is consistent with the common matrix operations. The same idea holds for writing down a block matrix representation for a linear operator with respect to a given subspace. Let $\mathcal{S} \subset \mathbb{C}^n$ be a k-dimensional vector subspace, such that $\mathbb{C}^n = \mathcal{S} \oplus \mathcal{S}^\perp$. Then, every linear operator in $\mathrm{End}(\mathbb{C}^n)$ can be rewritten as a 2×2-block-matrix, with its entries being matrices of a given dimension ie. [7]

$$
A \in \mathrm{End}(\mathbb{C}^n) \to \exists! \left\{ \begin{matrix} A_{\mathcal{S}} \in \mathrm{GL}(k, \mathbb{C}), & A_{\mathcal{S},\mathcal{S}^\perp} \in \mathrm{GL}(k \to n-k, \mathbb{C}), \\ A_{\mathcal{S}^\perp,\mathcal{S}} \in \mathrm{GL}(n-k \to k, \mathbb{C}), & A_{\mathcal{S}^\perp} \in \mathrm{GL}(n-k, \mathbb{C}), \end{matrix} \right\} /
$$

$$
A = \begin{pmatrix} A_{\mathcal{S}} & A_{\mathcal{S},\mathcal{S}^\perp} \\ A_{\mathcal{S}^\perp,\mathcal{S}} & A_{\mathcal{S}^\perp} \end{pmatrix} \begin{matrix} \mathcal{S} \\ \mathcal{S}^\perp \end{matrix}
$$

The Lax operator. Consider an XXX spin chain with N sites and a corresponding Hilbert space given by

(71)
$$
\mathbb{H} \simeq \bigotimes_{n=1}^{N} \mathbb{V}_n \quad \text{where } \mathbb{V}_n \simeq \mathbb{C}^2
$$
To these individual Hilbert spaces, additional auxiliary, non-physical, spaces $\mathbb{V}_a \simeq \mathbb{C}^2$, can be added.

The algebraic Bethe ansatz' basic tool is the lax operator $\mathcal{L}$, whose definition is as follows.

Definition 1. *The lax operator $\mathcal{L}$ can be defined as an operator which involves the local quantum space $\mathbb{V}_n$ and the auxiliary space $\mathbb{V}_a$, as follows*

---

[7]For example, consider the orthogonal projector over $\mathcal{S}$, labelled $P_{\mathcal{S}}$. In this notation, its matrix representation is given by

$$
P_{\mathcal{S}} = \begin{pmatrix} \mathcal{I} & 0 \\ 0 & 0 \end{pmatrix} \begin{matrix} \mathcal{S} \\ \mathcal{S}^\perp \end{matrix}
$$

where $\mathbb{1}_n$ and $\mathbf{S}_n^\alpha$ act on $\mathbb{V}_n$

$$\mathcal{L} : \mathbb{V}_n \otimes \mathbb{V}_a \to \mathbb{V}_n \otimes \mathbb{V}_a$$

(72)

$$\mathcal{L}_{n,a} = u(\mathbb{1}_n \otimes \mathbb{1}_a) + i \sum_\alpha \mathbf{S}_n^\alpha \otimes \sigma_a^\alpha, \ \forall n \leq N$$

while $\mathbb{1}_a$ and $\mathbf{S}_a^\alpha$ act upon $\mathbb{V}_a$ and where the

complex-valued u-parameter is the spectral parameter.

Note that equation (72) can be rewritten as

(73)

$$\mathcal{L}_{n,a} = u(\mathbb{1}_n \otimes \mathbb{1}_a) + i \sum_\alpha \mathbf{S}_n^\alpha \otimes \sigma_a^\alpha$$

$$= \left(\begin{array}{cc} u\mathbb{1}_n + i\mathbf{S}_n^z & i\mathbf{S}_n^x + \mathbf{S}_n^y \\ i\mathbf{S}_n^x - \mathbf{S}_n^y & u\mathbb{1}_n - i\mathbf{S}_n^z \end{array} \right)_a = \left(\begin{array}{cc} u\mathbb{1}_n + i\mathbf{S}_n^z & i\mathbf{S}_n^- \\ i\mathbf{S}_n^+ & u\mathbb{1}_n - i\mathbf{S}_n^z \end{array} \right)_a ,$$

which is a 2×2-matrix in the auxiliary space $\mathbb{V}_a$ and the matrix entries are operators acting on the physical Hilbert space $\mathbb{V}_n$. Consider now the following mathematical object.

Definition 2. *Let the permutation operator be*

(74)

$$\mathcal{P}_{n,a} = \frac{1}{2}\left(\mathbb{1}_n \otimes \mathbb{1}_a + \sum_\alpha \sigma_n^\alpha \otimes \sigma_n^\alpha \right) = \frac{1}{2}\sigma^{n\mu}\sigma_\mu^\alpha$$

which [8] acts on both the physical and auxiliary systems.

Using this definition for the permutation operator, the lax operator can then be rewritten as

(76)

$$\mathcal{L}_{n,a} = u(\mathbb{1}_n \otimes \mathbb{1}_a) + i\sum_\alpha \mathbf{S}_n^\alpha \otimes \sigma_a^\alpha = u(\mathbb{1}_n \otimes \mathbb{1}_a) + \frac{i}{2}\sum_\alpha \sigma_n^\alpha \otimes \sigma_a^\alpha$$

$$= u(\mathbb{1}_n \otimes \mathbb{1}_a) - \frac{i}{2}(\mathbb{1}_n \otimes \mathbb{1}_a) + \frac{i}{2}(\mathbb{1}_n \otimes \mathbb{1}_a) + \frac{i}{2}\sum_\alpha \sigma_n^\alpha \otimes \sigma_a^\alpha$$

$$= \left(u - \frac{i}{2} \right)(\mathbb{1}_n \otimes \mathbb{1}_a) + \frac{i}{2}\left(\mathbb{1}_n \otimes \mathbb{1}_a + \sum_\alpha \sigma_n^\alpha \otimes \sigma_n^\alpha \right)$$

$$= \left(u - \frac{i}{2} \right)\mathbb{1}_{n,a} + i\mathcal{P}_{n,a}.$$

Fundamental commutation Relations and Monodromy Matrix. Consider the following two lax operators

$$\mathcal{L}_{n,a_1}(u_1) : \ \mathbb{V}_n \otimes \mathbb{V}_{a_1} \to \mathbb{V}_n \otimes \mathbb{V}_{a_1}$$

$$\mathcal{L}_{n,a_2}(u_2) : \ \mathbb{V}_n \otimes \mathbb{V}_{a_2} \to \mathbb{V}_n \otimes \mathbb{V}_{a_2}$$

both of which act on the physical quantum state and on two different auxiliary spaces as well[9].

The product of these two operator is then a triple tensor product in $\mathbb{V}_n \otimes \mathbb{V}_{a_1} \otimes \mathbb{V}_{a_2}$. Now, the following matrix arises

---

[8]Note that the permutation operator can be written as a 2×2-matrix in the auxiliary space $\mathbb{V}_a$,

(75)

$$\mathcal{P}_{n,a} = \frac{1}{2}\left[\left(\begin{array}{cc} \mathbb{1}_n & 0 \\ 0 & \mathbb{1}_a \end{array} \right) + \sigma_n^x \otimes \left(\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right)_a + \sigma_n^y \otimes \left(\begin{array}{cc} 0 & -i \\ i & 0 \end{array} \right)_a + \sigma_n^z \otimes \left(\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array} \right)_a \right]$$

$$= \left(\begin{array}{cc} \mathbb{1}_n + \sigma_n^z & \sigma_n^x - i\sigma_n^y \\ \sigma_n^x + i\sigma_n^y & \mathbb{1}_n - \sigma_n^z \end{array} \right)_a = \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right)$$

Definition 3. *Consider an $\mathcal{R}$-matrix which relates the permutation and spectral parameters of two auxiliary spaces, which is given by*

(78)
$$\mathcal{R}: \ \mathrm{V}_{a_1} \otimes \mathrm{V}_{a_2} \otimes \mathrm{V}_{a_1} \otimes \mathrm{V}_{a_2}$$

$$\mathcal{R}_{a_1,a_2}(u_1 - u_2) = (u_1 - u_2)\mathbb{1}_{a_1,a_2} + i\mathcal{P}_{a_1,a_2},$$

which[10] doesn't act on the physical system at all, but on the auxiliary, phantom, systems.

An interesting relationship between the products of the two lax operators and the $\mathcal{R}$-matrix can then be found[11], consider

$$\mathcal{R}_{a_1,a_2}(u_1 - u_2)\mathcal{L}_{n,a_1}(u_1)\mathcal{L}_{n,a_2}(u_2)$$

$$= \left((u_1 - u_2)\mathbb{1}_{a_1,a_2} + i\mathcal{P}_{a_1,a_2}\right)\left(\left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_1} + i\mathcal{P}_{n,a_1}\right)\left(\left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_2} + i\mathcal{P}_{n,a_2}\right)$$

$$= \left((u_1 - u_2)\mathbb{1}_{a_1,a_2}\left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_1} + (u_1 - u_2)\mathbb{1}_{a_1,a_2}i\mathcal{P}_{n,a_1} + i\mathcal{P}_{a_1,a_2}\left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_1} + i^2\mathcal{P}_{a_1,a_2}\mathcal{P}_{n,a_1}\right)$$

$$\times \left(\left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_2} + i\mathcal{P}_{n,a_2}\right)$$

$$= (u_1 - u_2)\mathbb{1}_{a_1,a_2}\left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_1}\left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_2} + (u_1 - u_2)\mathbb{1}_{a_1,a_2}\left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_1}i\mathcal{P}_{n,a_2}$$

$$+ (u_1 - u_2)\mathbb{1}_{a_1,a_2}i\mathcal{P}_{n,a_1}\left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_2} + (u_1 - u_2)\mathbb{1}_{a_1,a_2}i\mathcal{P}_{n,a_1}i\mathcal{P}_{n,a_2}$$

$$+ i\mathcal{P}_{a_1,a_2}\left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_1}\left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_2} + i\mathcal{P}_{a_1,a_2}\left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_1}i\mathcal{P}_{n,a_2}$$

$$- \mathcal{P}_{a_1,a_2}\mathcal{P}_{n,a_1}\left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_2} - \mathcal{P}_{a_1,a_2}\mathcal{P}_{n,a_1}i\mathcal{P}_{n,a_2}$$

$$= (u_1 - u_2)\mathbb{1}_{a_1,a_2}\left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_2} \times \left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_1} + (u_1 - u_2)\mathbb{1}_{a_1,a_2}i\mathcal{P}_{n,a_2} \times \left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_1}$$

$$+ (u_1 - u_2)\mathbb{1}_{a_1,a_2}\left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_2} \times i\mathcal{P}_{n,a_1} + (u_1 - u_2)\mathbb{1}_{a_1,a_2}i\mathcal{P}_{n,a_2} \times i\mathcal{P}_{n,a_1}$$

$$+ i\mathcal{P}_{a_1,a_2}\left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_1}\left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_2} + i\mathcal{P}_{a_1,a_2}\left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_1}i\mathcal{P}_{n,a_2}$$

$$- \mathcal{P}_{a_1,a_2}\mathcal{P}_{n,a_1}\left(u - \frac{i}{2}\right)\mathbb{1}_{n,a_2} - \mathcal{P}_{a_1,a_2}\mathcal{P}_{n,a_1}i\mathcal{P}_{n,a_2}$$

---

[10]Note that $\mathcal{R}$ can be rewritten as a 2×2-matrix in either of the auxiliary spaces, as follows

(79) $\quad \mathcal{R}_{a1,a2} = \begin{pmatrix} \left(u + \frac{i}{2}\right)\mathbb{1}_{a_2} + i\mathbf{S}^z_{a_2} & i\mathbf{S}^-_{a_2} \\ i\mathbf{S}^+_{a_2} & \left(u + \frac{i}{2}\right)\mathbb{1}_{a_2} - i\mathbf{S}^z_{a_2} \end{pmatrix}_{a_1} = \begin{pmatrix} \left(u + \frac{i}{2}\right)\mathbb{1}_{a_1} + i\mathbf{S}^z_{a_1} & i\mathbf{S}^-_{a_1} \\ i\mathbf{S}^+_{a_1} & \left(u + \frac{i}{2}\right)\mathbb{1}_{a_1} - i\mathbf{S}^z_{a_2} \end{pmatrix}_{a_2}.$

[11]Note that the following permutation properties hold

(80)
$$\mathcal{P}_{n,a_1}\mathcal{P}_{n,a_2} = \mathcal{P}_{a_1,a_2}\mathcal{P}_{n,a_1} = \mathcal{P}_{n,a_2}\mathcal{P}_{a_2,a_1} \text{ and } \mathcal{P}_{a,b} = \mathcal{P}_{b,a}$$

$$= \left((u_1 - u_2)\mathbb{1}_{a_1,a_2}\left(u - \frac{i}{2} \right)\mathbb{1}_{n,a_2} + (u_1 - u_2)\mathbb{1}_{a_1,a_2}i\boldsymbol{\mathcal{P}}_{n,a_2} \right)\left(u - \frac{i}{2} \right)\mathbb{1}_{n,a_1}$$

$$+ \left((u_1 - u_2)\mathbb{1}_{a_1,a_2}\left(u - \frac{i}{2} \right)\mathbb{1}_{n,a_2} + (u_1 - u_2)\mathbb{1}_{a_1,a_2}i\boldsymbol{\mathcal{P}}_{n,a_2} \right) \times i\boldsymbol{\mathcal{P}}_{n,a_1}$$

$$+ i\boldsymbol{\mathcal{P}}_{a_1,a_2}\left(u - \frac{i}{2} \right)\mathbb{1}_{n,a_1}\left(u - \frac{i}{2} \right)\mathbb{1}_{n,a_2} + i\boldsymbol{\mathcal{P}}_{a_1,a_2}\left(u - \frac{i}{2} \right)\mathbb{1}_{n,a_1}i\boldsymbol{\mathcal{P}}_{n,a_2}$$

$$- \boldsymbol{\mathcal{P}}_{a_1,a_2}\boldsymbol{\mathcal{P}}_{n,a_1}\left(u - \frac{i}{2} \right)\mathbb{1}_{n,a_2} - \boldsymbol{\mathcal{P}}_{a_1,a_2}\boldsymbol{\mathcal{P}}_{n,a_1}i\boldsymbol{\mathcal{P}}_{n,a_2}$$

$$= \left((u_1 - u_2)\mathbb{1}_{a_1,a_2}\left(u - \frac{i}{2} \right)\mathbb{1}_{n,a_2} + (u_1 - u_2)\mathbb{1}_{a_1,a_2}i\boldsymbol{\mathcal{P}}_{n,a_2} \right)\left(\left(u - \frac{i}{2} \right)\mathbb{1}_{n,a_2} + i\boldsymbol{\mathcal{P}}_{n,a_1} \right)$$

$$+ i\boldsymbol{\mathcal{P}}_{a_1,a_2}\left(u - \frac{i}{2} \right)\mathbb{1}_{n,a_1}\left(u - \frac{i}{2} \right)\mathbb{1}_{n,a_2} + i\boldsymbol{\mathcal{P}}_{a_1,a_2}\left(u - \frac{i}{2} \right)\mathbb{1}_{n,a_1}i\boldsymbol{\mathcal{P}}_{n,a_2}$$

$$- \boldsymbol{\mathcal{P}}_{a_1,a_2}\boldsymbol{\mathcal{P}}_{n,a_1}\left(u - \frac{i}{2} \right)\mathbb{1}_{n,a_2} - \boldsymbol{\mathcal{P}}_{a_1,a_2}\boldsymbol{\mathcal{P}}_{n,a_1}i\boldsymbol{\mathcal{P}}_{n,a_2}$$

$$= \left(\left(u - \frac{i}{2} \right)\mathbb{1}_{n,a_2} + i\boldsymbol{\mathcal{P}}_{n,a_2} \right)\left(\left(u - \frac{i}{2} \right)\mathbb{1}_{n,a_2} + i\boldsymbol{\mathcal{P}}_{n,a_1} \right)(u_1 - u_2)\mathbb{1}_{a_1,a_2}$$

$$+ \left(\left(u - \frac{i}{2} \right)\mathbb{1}_{n,a_2} + i\boldsymbol{\mathcal{P}}_{n,a_2} \right)\left(\left(u - \frac{i}{2} \right)\mathbb{1}_{n,a_2} + i\boldsymbol{\mathcal{P}}_{n,a_1} \right)i\boldsymbol{\mathcal{P}}_{n,a_2}$$

$$= \left(\left(u - \frac{i}{2} \right)\mathbb{1}_{n,a_2} + i\boldsymbol{\mathcal{P}}_{n,a_2} \right)\left(\left(u - \frac{i}{2} \right)\mathbb{1}_{n,a_2} + i\boldsymbol{\mathcal{P}}_{n,a_1} \right)\left(\left(u - \frac{i}{2} \right)\mathbb{1}_{n,a_2} + i\boldsymbol{\mathcal{P}}_{n,a_2} \right),$$

which proves that

$$(81) \qquad \boxed{\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1 - u_2)\boldsymbol{\mathcal{L}}_{n,a_1}(u_1)\boldsymbol{\mathcal{L}}_{n,a_2}(u_2) = \boldsymbol{\mathcal{L}}_{n,a_1}(u_2)\boldsymbol{\mathcal{L}}_{n,a_1}(u_1)\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1 - u_2)}.$$

From the fundamental commutation relation it can then be shown that the $\boldsymbol{\mathcal{R}}$-matrix follows the quantum Yang-Baxter equation. Consider then the product of the following $\boldsymbol{\mathcal{L}}$-operators,

$$(82) \qquad \begin{aligned} \boldsymbol{\mathcal{L}}_{n,1}\boldsymbol{\mathcal{L}}_{n,2}\boldsymbol{\mathcal{L}}_{n,3} &= \boldsymbol{\mathcal{R}}_{12}^{-1}\boldsymbol{\mathcal{L}}_{n,2}\boldsymbol{\mathcal{L}}_{n,1}\boldsymbol{\mathcal{L}}_{n,3}\boldsymbol{\mathcal{R}}_{12} \\ &= \boldsymbol{\mathcal{R}}_{12}^{-1}\boldsymbol{\mathcal{R}}_{13}^{-1}\boldsymbol{\mathcal{L}}_{n,2}\boldsymbol{\mathcal{L}}_{n,3}\boldsymbol{\mathcal{L}}_{n,1}\boldsymbol{\mathcal{R}}_{13}\boldsymbol{\mathcal{R}}_{12} \\ &= \boldsymbol{\mathcal{R}}_{12}^{-1}\boldsymbol{\mathcal{R}}_{13}^{-1}\boldsymbol{\mathcal{R}}_{23}^{-1}\boldsymbol{\mathcal{L}}_{n,3}\boldsymbol{\mathcal{L}}_{n,2}\boldsymbol{\mathcal{L}}_{n,1}\boldsymbol{\mathcal{R}}_{23}\boldsymbol{\mathcal{R}}_{13}\boldsymbol{\mathcal{R}}_{12} \\ &= (\boldsymbol{\mathcal{R}}_{23}\boldsymbol{\mathcal{R}}_{13}\boldsymbol{\mathcal{R}}_{12})^{-1}\boldsymbol{\mathcal{L}}_{n,3}\boldsymbol{\mathcal{L}}_{n,2}\boldsymbol{\mathcal{L}}_{n,1}\boldsymbol{\mathcal{R}}_{23}\boldsymbol{\mathcal{R}}_{13}\boldsymbol{\mathcal{R}}_{12}, \end{aligned}$$

and similarly

$$(83) \qquad \begin{aligned} \boldsymbol{\mathcal{L}}_{n,1}\boldsymbol{\mathcal{L}}_{n,2}\boldsymbol{\mathcal{L}}_{n,3} &= \boldsymbol{\mathcal{R}}_{23}^{-1}\boldsymbol{\mathcal{L}}_{n,1}\boldsymbol{\mathcal{L}}_{n,3}\boldsymbol{\mathcal{L}}_{n,2}\boldsymbol{\mathcal{R}}_{23} \\ &= \boldsymbol{\mathcal{R}}_{23}^{-1}\boldsymbol{\mathcal{R}}_{13}^{-1}\boldsymbol{\mathcal{L}}_{n,3}\boldsymbol{\mathcal{L}}_{n,1}\boldsymbol{\mathcal{L}}_{n,2}\boldsymbol{\mathcal{R}}_{13}\boldsymbol{\mathcal{R}}_{23} \\ &= \boldsymbol{\mathcal{R}}_{23}^{-1}\boldsymbol{\mathcal{R}}_{13}^{-1}\boldsymbol{\mathcal{R}}_{12}^{-1}\boldsymbol{\mathcal{L}}_{n,3}\boldsymbol{\mathcal{L}}_{n,2}\boldsymbol{\mathcal{L}}_{n,1}\boldsymbol{\mathcal{R}}_{12}\boldsymbol{\mathcal{R}}_{13}\boldsymbol{\mathcal{R}}_{23} \\ &= (\boldsymbol{\mathcal{R}}_{12}\boldsymbol{\mathcal{R}}_{13}\boldsymbol{\mathcal{R}}_{23})^{-1}\boldsymbol{\mathcal{L}}_{n,3}\boldsymbol{\mathcal{L}}_{n,2}\boldsymbol{\mathcal{L}}_{n,1}\boldsymbol{\mathcal{R}}_{12}\boldsymbol{\mathcal{R}}_{13}\boldsymbol{\mathcal{R}}_{23}, \end{aligned}$$

both of which then yield an important constraint on the $\boldsymbol{\mathcal{R}}$-matrix, namely

$$(84) \qquad \boxed{\boldsymbol{\mathcal{R}}_{12}\boldsymbol{\mathcal{R}}_{13}\boldsymbol{\mathcal{R}}_{23} = \boldsymbol{\mathcal{R}}_{23}\boldsymbol{\mathcal{R}}_{13}\boldsymbol{\mathcal{R}}_{12}},$$

which is the quantum Yang-Baxter equation.

Definition 4. *Let the monodromy matrix $\boldsymbol{\mathcal{T}}_{N,a}$ be defined as a product of subsequent of $\boldsymbol{\mathcal{L}}$-operators, ie.*

$$\boldsymbol{\mathcal{T}}_{N,a} = \prod_1^{j=N} \boldsymbol{\mathcal{L}}_{j,a}$$

(85)

$$\boldsymbol{\mathcal{T}}_{N,a} = \prod_1^{j=N} \begin{pmatrix} u\mathbb{1}_j + i\mathbf{S}_j^z & i\mathbf{S}_j^- \\ i\mathbf{S}_j^+ & u\mathbb{1}_j - i\mathbf{S}_j^z \end{pmatrix} \equiv \begin{pmatrix} A(u) & B(u) \\ C(u) & D(u) \end{pmatrix},$$

This matrix obeys the fundamental relation as well[12], ie.

(87)

$$\boxed{\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1 - u_2)\boldsymbol{\mathcal{T}}_{N,a_1}(u_1)\boldsymbol{\mathcal{T}}_{N,a_2}(u_2) = \boldsymbol{\mathcal{T}}_{N,a_2}(u_2)\boldsymbol{\mathcal{T}}_{N,a_1}(u_1)\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1 - u_2)}.$$

Definition 5. *The transfer $\mathfrak{t}$-matrix is defined as the $\boldsymbol{\mathcal{T}}$-matrix's partial trace over the auxiliary space $\mathbb{V}_a$, ie.*

(88)

$$\mathfrak{t}(u) = \mathrm{Tr}_{\mathbb{V}_a} \boldsymbol{\mathcal{T}}_{N,a}(u) = \mathrm{Tr}_{\mathbb{V}_a} \begin{pmatrix} A(u) & B(u) \\ C(u) & D(u) \end{pmatrix} = A(u) + D(u),$$

where $\mathfrak{t}, A, B, C, D$ all are 2×2-matrices in the physical system, the auxiliary systems being present no more.

With the previous definitions in mind, equation (87)'s double trace can then be written in terms of the transfer matrix, as follows

$$\mathrm{Tr}_{\mathbb{V}_{a_1}} \mathrm{Tr}_{\mathbb{V}_{a_2}} \left[\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1 - u_2)\boldsymbol{\mathcal{T}}_{N,a_1}(u_1)\boldsymbol{\mathcal{T}}_{N,a_2}(u_2) \right] =$$

$$\Rightarrow \mathrm{Tr}_{\mathbb{V}_{a_1}} \mathrm{Tr}_{\mathbb{V}_{a_2}} \left[\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1 - u_2)\boldsymbol{\mathcal{T}}_{N,a_1}(u_1)\boldsymbol{\mathcal{T}}_{N,a_2}(u_2) - \boldsymbol{\mathcal{T}}_{N,a_1}(u_2)\boldsymbol{\mathcal{T}}_{N,a_1}(u_1)\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1 - u_2) \right] = 0$$

$$\Rightarrow \mathrm{Tr}_{\mathbb{V}_{a_1}} \mathrm{Tr}_{\mathbb{V}_{a_2}} \left[\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1 - u_2)\left(\boldsymbol{\mathcal{T}}_{N,a_1}(u_1)\boldsymbol{\mathcal{T}}_{N,a_2}(u_2) - \boldsymbol{\mathcal{T}}_{N,a_2}(u_2)\boldsymbol{\mathcal{T}}_{N,a_1}(u_1) \right) \right] = 0$$

which implies $\mathfrak{t}(u_1)\mathfrak{t}(u_2) = \mathfrak{t}(u_2)\mathfrak{t}(u_1)$.

Therefore the transfer matrix commutes with itself for different values of the spectral parameter. It turns out that the monodromy matrix can be expanded in a power series around any point $z_0 \in \mathbb{C}$, generating an infinite set of linearly independent commuting operators acting on the full quantum space. This formally leads to the model's integrability.

---

[12]For example, consider $N = 2$, then

$$\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1 - u_2)\boldsymbol{\mathcal{T}}_{N,a_1}(u_1)\boldsymbol{\mathcal{T}}_{N,a_2}(u_2) = \boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1 - u_2)\boldsymbol{\mathcal{T}}_{2,a_1}(u_1)\boldsymbol{\mathcal{T}}_{2,a_2}(u_2)$$

$$= \boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1 - u_2)\boldsymbol{\mathcal{L}}_{2,a_1}(u_1)\boldsymbol{\mathcal{L}}_{2,a_1}(u_1)\boldsymbol{\mathcal{L}}_{2,a_2}(u_2)\boldsymbol{\mathcal{L}}_{1,a_2}(u_2)$$

$$= \left(\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1 - u_2)\boldsymbol{\mathcal{L}}_{2,a_1}(u_1)\boldsymbol{\mathcal{L}}_{2,a_2}(u_2) \right)\boldsymbol{\mathcal{L}}_{1,a_1}(u_1)\boldsymbol{\mathcal{L}}_{1,a_2}(u_2)$$

(86)

$$= \boldsymbol{\mathcal{L}}_{2,a_2}(u_2)\boldsymbol{\mathcal{L}}_{2,a_1}(u_1)\left(\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1 - u_2)\boldsymbol{\mathcal{L}}_{1,a_1}(u_1)\boldsymbol{\mathcal{L}}_{1,a_2}(u_2) \right)$$

$$= \boldsymbol{\mathcal{L}}_{2,a_2}(u_2)\boldsymbol{\mathcal{L}}_{2,a_1}(u_1)\boldsymbol{\mathcal{L}}_{1,a_2}(u_2)\boldsymbol{\mathcal{L}}_{1,a_1}(u_1)\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1 - u_2)$$

$$= \boldsymbol{\mathcal{L}}_{2,a_2}(u_2)\boldsymbol{\mathcal{L}}_{1,a_2}(u_2)\boldsymbol{\mathcal{L}}_{2,a_1}(u_1)\boldsymbol{\mathcal{L}}_{1,a_1}(u_1)\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1 - u_2)$$

$$= \boldsymbol{\mathcal{T}}_{n,a_1}(u_2)\boldsymbol{\mathcal{T}}_{n,a_1}(u_1)\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1 - u_2),$$

where the fundamental commutation relation given by (81) was used. Note as well that $\boldsymbol{\mathcal{L}}$-operators acting on different spaces commute as well, eg. $[\boldsymbol{\mathcal{L}}_{n_1,a_1}, \boldsymbol{\mathcal{L}}_{n_2,a_2}] \propto \delta_{n_1,n_2}$. The general fundamental relation can then be proved via mathematical induction for the general case.

Monodromy matrix and the Hamiltonian. The Hamiltonian operator belongs to the family of transfer matrices. Let $z_0 = \frac{i}{2}$, then the monodromy matrix can be expanded around said complex point yielding (up to first order)

$$
\begin{aligned}
\boldsymbol{\mathcal{T}}_{N,a}(z_0) &= \prod_{1}^{j=N} \left[\left(z_0 - \frac{i}{2} \right) \mathbb{1}_{j,a} + i \boldsymbol{\mathcal{P}}_{j,a} \right] = i^N \boldsymbol{\mathcal{P}}_{N,a} \boldsymbol{\mathcal{P}}_{N-1,a} \cdots \boldsymbol{\mathcal{P}}_{1,a} = \\
&= i^N \boldsymbol{\mathcal{P}}_{1,2} \boldsymbol{\mathcal{P}}_{2,3} \cdots \boldsymbol{\mathcal{P}}_{N-1,N} \boldsymbol{\mathcal{P}}_{N,a} \\
&= i^N \prod_{j=1}^{N-1} \boldsymbol{\mathcal{P}}_{j,\,j+1} \boldsymbol{\mathcal{P}}_{N,a}
\end{aligned}
\tag{89}
$$

where the second line holds given the properties of the permutation operators. Taking the partial trace over the auxiliary space yields the transfer matrix, as follows

$$
\begin{aligned}
\mathsf{t}(z_0) &= \mathrm{Tr}_{\mathbb{V}_a} \boldsymbol{\mathcal{T}}_{N,a}(z_0) = \mathrm{Tr}_{\mathbb{V}_a} i^N \boldsymbol{\mathcal{P}}_{1,2} i^N \prod_{j=1}^{N-1} \boldsymbol{\mathcal{P}}_{j,\,j+1} \boldsymbol{\mathcal{P}}_{N,a} \\
&= i^N \prod_{j=1}^{N-1} \boldsymbol{\mathcal{P}}_{j,\,j+1} \mathrm{Tr}_{\mathbb{V}_a} \boldsymbol{\mathcal{P}}_{N,a} = i^N \prod_{j=1}^{N-1} \boldsymbol{\mathcal{P}}_{j,\,j+1} \mathrm{Tr}_{\mathbb{V}_a} \frac{1}{2} \begin{pmatrix} \mathbb{1}_n + \sigma_n^z & \sigma_n^x - i\sigma_n^y \\ \sigma_n^x + i\sigma_n^y & \mathbb{1}_n - \sigma_n^z \end{pmatrix}_a. \\
&= i^N \prod_{j=1}^{N-1} \boldsymbol{\mathcal{P}}_{j,\,j+1} \mathbb{1}_N.
\end{aligned}
\tag{90}
$$

Then $\mathcal{U} = i^{-N} \mathsf{t}(z_0) = \prod_{j=1}^{N-1} \boldsymbol{\mathcal{P}}_{j,\,j+1}$ is a shift operator in the full quantum Hilbert space $\mathbb{H}$. In effect, note that

$$
\boldsymbol{\mathcal{P}}_{n_1,\,n_2} \mathbf{X}_{n_2} \boldsymbol{\mathcal{P}}_{n_1,\,n_2} = \mathbf{X}_{n_1},
$$

ie. this permutation moves $\mathbf{X}$ one step back. Furthermore, this operator is unitary (given that the permutations are, by definition, unitary operators). Then, according to Stone's theorem on one-parameter unitary groups, the shift operator is related to the momentum operator

$$
\mathcal{U} = e^{iP}
\tag{91}
$$

The next order in the expansion of the transfer matrix can then be found via the derivative of the monodromy matrix at $z_0 = \frac{i}{2}$,

$$\frac{d\boldsymbol{\mathcal{T}}_{N,a}(u)}{du}\Big|_{u=z_0} = \frac{d}{du}\prod_{1}^{j=N}\begin{pmatrix} u\mathbb{1}_j + i\mathbf{S}_j^z & i\mathbf{S}_j^- \\ i\mathbf{S}_j^+ & u\mathbb{1}_j - i\mathbf{S}_j^z \end{pmatrix}\Big|_{u=z_0}$$

$$= \frac{d}{du}\left[\begin{pmatrix} u\mathbb{1}_N + i\mathbf{S}_N^z & i\mathbf{S}_N^- \\ i\mathbf{S}_N^+ & u\mathbb{1}_N - i\mathbf{S}_N^z \end{pmatrix}\right]\Big|_{u=z_0}\prod_{1}^{j=N-1}\begin{pmatrix} u\mathbb{1}_j + i\mathbf{S}_j^z & i\mathbf{S}_j^- \\ i\mathbf{S}_j^+ & u\mathbb{1}_j - i\mathbf{S}_j^z \end{pmatrix}$$

$$+ \begin{pmatrix} u\mathbb{1}_N + i\mathbf{S}_N^z & i\mathbf{S}_N^- \\ i\mathbf{S}_N^+ & u\mathbb{1}_N - i\mathbf{S}_N^z \end{pmatrix}\frac{d}{du}\left[\begin{pmatrix} u\mathbb{1}_{N-1} + i\mathbf{S}_{N-1}^z & i\mathbf{S}_{N-1}^- \\ i\mathbf{S}_{N-1}^+ & u\mathbb{1}_{N-1} - i\mathbf{S}_{N-1}^z \end{pmatrix}\right]\Big|_{u=z_0}$$

$$\times \prod_{1}^{j=N-2}\begin{pmatrix} u\mathbb{1}_j + i\mathbf{S}_j^z & i\mathbf{S}_j^- \\ i\mathbf{S}_j^+ & u\mathbb{1}_j - i\mathbf{S}_j^z \end{pmatrix}$$

$$+ \cdots + \prod_{3}^{j=N}\begin{pmatrix} u\mathbb{1}_j + i\mathbf{S}_j^z & i\mathbf{S}_j^- \\ i\mathbf{S}_j^+ & u\mathbb{1}_j - i\mathbf{S}_j^z \end{pmatrix}\frac{d}{du}\left[\begin{pmatrix} u\mathbb{1}_2 + i\mathbf{S}_2^z & i\mathbf{S}_2^- \\ i\mathbf{S}_2^+ & u\mathbb{1}_2 - i\mathbf{S}_2^z \end{pmatrix}\right]\Big|_{u=z_0}$$

$$\times \begin{pmatrix} u\mathbb{1}_1 + i\mathbf{S}_1^z & i\mathbf{S}_1^- \\ i\mathbf{S}_1^+ & u\mathbb{1}_1 - i\mathbf{S}_1^z \end{pmatrix}$$

$$+ \prod_{2}^{j=N}\begin{pmatrix} u\mathbb{1}_j + i\mathbf{S}_j^z & i\mathbf{S}_j^- \\ i\mathbf{S}_j^+ & u\mathbb{1}_j - i\mathbf{S}_j^z \end{pmatrix}\frac{d}{du}\left[\begin{pmatrix} u\mathbb{1}_1 + i\mathbf{S}_1^z & i\mathbf{S}_1^- \\ i\mathbf{S}_1^+ & u\mathbb{1}_1 - i\mathbf{S}_1^z \end{pmatrix}\right]\Big|_{u=z_0}s$$

$$= i^{N-1}\Big[\prod_{1}^{j=N-1}\boldsymbol{\mathcal{P}}_j + \prod_{N}^{j=N}\boldsymbol{\mathcal{P}}_j \times \prod_{1}^{j=N-2}\boldsymbol{\mathcal{P}}_j + \cdots + \prod_{k}^{j=N}\boldsymbol{\mathcal{P}}_j \times \prod_{1}^{k+2}\boldsymbol{\mathcal{P}}_j$$

$$+ \cdots + \prod_{3}^{j=N}\boldsymbol{\mathcal{P}}_j \times \prod_{1}^{j=1}\boldsymbol{\mathcal{P}}_j + \prod_{2}^{j=N}\boldsymbol{\mathcal{P}}_j\Big]$$

$$= i^{N-1}\sum_{n\in\mathbb{N}}\boldsymbol{\mathcal{P}}_{N,a}\cdots\boldsymbol{\mathcal{P}}_{n-1,a}\boldsymbol{\mathcal{P}}_{n+1,a}\cdots\boldsymbol{\mathcal{P}}_{1,a}$$

$$= i^{N-1}\sum_{n\in\mathbb{N}}\boldsymbol{\mathcal{P}}_{1,2}\boldsymbol{\mathcal{P}}_{2,3}\cdots\boldsymbol{\mathcal{P}}_{n-1,n+1}\cdots\boldsymbol{\mathcal{P}}_{N-1,N}\boldsymbol{\mathcal{P}}_{N,a}$$

$$= i^{N-1}\sum_{n\in\mathbb{N}}\prod_{j=1}^{n-2}\boldsymbol{\mathcal{P}}_{j,\,j+1} \times \boldsymbol{\mathcal{P}}_{n-1,\,n+1} \times \prod_{j=n+2}^{N-1}\boldsymbol{\mathcal{P}}_{j,\,j+1}\boldsymbol{\mathcal{P}}_{N,a}.$$

Therefore,

$$\frac{d\mathfrak{t}(u)}{du}\Big|_{u=z_0} = \frac{d\,\mathrm{Tr}_{\mathbb{V}_a}\,\boldsymbol{\mathcal{T}}_{N,a}(u)}{du}\Big|_{u=z_0}$$

(92)
$$= i^{N-1}\sum_{n\in\mathbb{N}}\prod_{j=1}^{n-2}\boldsymbol{\mathcal{P}}_{j,\,j+1} \times \boldsymbol{\mathcal{P}}_{n-1,\,n+1} \times \prod_{j=n+2}^{N-1}\boldsymbol{\mathcal{P}}_{j,\,j+1} \times \mathrm{Tr}_{\mathbb{V}_a}\,\boldsymbol{\mathcal{P}}_{N,a}$$

$$= i^{N-1}\sum_{n\in\mathbb{N}}\prod_{j=1}^{n-2}\boldsymbol{\mathcal{P}}_{j,\,j+1} \times \boldsymbol{\mathcal{P}}_{n-1,\,n+1} \times \prod_{j=n+2}^{N-1}\boldsymbol{\mathcal{P}}_{j,\,j+1}.$$

Now, note that

$$
\begin{aligned}
\frac{d}{du}\log(\mathfrak{t}(u))\Big|_{u=z_0} &= \frac{d}{du}\log \mathrm{Tr}_{\mathbb{V}_a}\boldsymbol{\mathcal{T}}(u)\Big|_{u=z_0} \\
&= \frac{d}{du}\Big(\mathfrak{t}(u)\Big)\mathfrak{t}(u)^{-1}\Big|_{u=z_0} \\
&= i^{N-1}\sum_{n\in\mathbb{N}}\left(\prod_{j=1}^{n-2}\boldsymbol{\mathcal{P}}_{j,\,j+1}\times \boldsymbol{\mathcal{P}}_{n-1,\,n+1}\times \prod_{j=n+2}^{N-1}\boldsymbol{\mathcal{P}}_{j,\,j+1}\right)\times \left(i^N\prod_{j=1}^{N-1}\boldsymbol{\mathcal{P}}_{j,\,j+1}\mathbb{1}_N\right)^{-1} \\
&= \frac{1}{i}\sum_{n\in\mathbb{N}}\left(\prod_{j=1}^{n-2}\boldsymbol{\mathcal{P}}_{j,\,j+1}\times \boldsymbol{\mathcal{P}}_{n-1,\,n+1}\times \prod_{j=n+2}^{N-1}\boldsymbol{\mathcal{P}}_{j,\,j+1}\right)\times \prod_{j=1}^{N-1}\boldsymbol{\mathcal{P}}_{N-1-j,\,N-j}^{-1} \\
&= \frac{1}{i}\sum_{n\in\mathbb{N}}\left(\prod_{j=1}^{n-2}\boldsymbol{\mathcal{P}}_{j,\,j+1}\times \boldsymbol{\mathcal{P}}_{n-1,\,n+1}\times \prod_{j=n+2}^{N-1}\boldsymbol{\mathcal{P}}_{j,\,j+1}\right)\times \prod_{1}^{j=N-1}\boldsymbol{\mathcal{P}}_{j,j+1} \\
&= \frac{1}{i}\sum_{n\in\mathbb{N}}\left(\cancel{\boldsymbol{\mathcal{P}}_{1,2}}\cancel{\boldsymbol{\mathcal{P}}_{2,3}}\cdots \boldsymbol{\mathcal{P}}_{n-1,\,n+1}\cdots \cancel{\boldsymbol{\mathcal{P}}_{N-1,N}}\times \cancel{\boldsymbol{\mathcal{P}}_{N-1,\,N}}\cdots \boldsymbol{\mathcal{P}}_{n,\,n+1}\cdots \cancel{\boldsymbol{\mathcal{P}}_{2,3}}\cancel{\boldsymbol{\mathcal{P}}_{1,2}}\right) \\
&= \frac{1}{i}\sum_{n\in\mathbb{N}}\boldsymbol{\mathcal{P}}_{N-1,\,N}
\end{aligned}
\tag{93}
$$

Note that (69) can be rewritten in terms of the permutation operator as

$$
\mathbf{H} = \frac{J}{2}\sum_n\left(\boldsymbol{\mathcal{P}}_{n,\,n+1}-\mathbb{1}^{\otimes N}\right),
\tag{94}
$$

which in turn can be rewritten as

$$
\boxed{\mathbf{H} = \frac{J}{2}\left(i\frac{d}{du}\log(\mathfrak{t}(u))\Big|_{u=z_0}-N\mathbb{1}^{\otimes N}\right),}
\tag{95}
$$

which shows that the XXX Hamiltonian belongs to the family of $N-1$ commuting operators generated by the trace of the monodromy matrix $\boldsymbol{\mathcal{T}}$. As a result, the Hamiltonian commutes with the transfer matrix $[\mathbf{H},\mathfrak{t}(u)]=0$.

Diagonalizing the Hamiltonian. The only task left is to diagonalize the Hamiltonian, by diagonalizing the transfer matrix. The monodromy matrix $\mathfrak{t}_{N,a}$ is a 2×2-matrix in the physical space $\mathbb{V}_a$. Recalling the fundamental commutation relation between the $\mathcal{R}$ and $\mathcal{T}$-matrices, then the following theorem holds,

Theorem 1. *Let $\boldsymbol{\mathcal{T}}$ be the monodromy matrix for the Heisenberg model. Then, $\forall u_1, u_2 \in \mathbb{C}$ then*

- $[\boldsymbol{\mathcal{T}}_{ij}(u_1),\boldsymbol{\mathcal{T}}_{ij}(u_2)]=0$, $i,j=1,2$, *which entails that* $[B(u_1),B(u_2)]=\cdots=[D(u_1),D(u_2)]=0$

- $A(u_1)B(u_2)=f(u_1-u_2)B(u_2)A(u_1)+g(u_1-u_2)B(u_1)A(u_2)$

- $D(u_1)B(u_2)=h(u_1-u_2)B(u_2)D(u_1)+k(u_1-u_2)B(u_1)D(u_2)$,

where f, g, h, and k are given by

$$
f(u)=\frac{u-i}{u}, \qquad g(u)=\frac{i}{u}, \qquad h(u)=\frac{u+i}{u}, \qquad k(u)=-\frac{i}{u}.
$$

Proof. Consider two different monodromy matrices, labelled $\boldsymbol{\mathcal{T}}_{N,a_1}(u_1)$ and $\boldsymbol{\mathcal{T}}_{N,a_2}(u_2)$, acting on two different auxiliary spaces, labelled $\mathbb{V}_1$ and $\mathbb{V}_2$ respectively. Then, these $\boldsymbol{\mathcal{T}}$-matrices can be written as

$$\boldsymbol{\mathcal{T}}_{N,a_1}(u_1) = \begin{pmatrix} A(u_1) & B(u_1) \\ C(u_1) & D(u_1) \end{pmatrix} \otimes \mathbb{1}_{V_2}, \quad \boldsymbol{\mathcal{T}}_{N,a_2}(u_2) = \mathbb{1}_{V_1} \otimes \begin{pmatrix} A(u_2) & B(u_2) \\ C(u_2) & D(u_2) \end{pmatrix}$$

$$(96) \qquad = \begin{pmatrix} A(u_1) & 0 & B(u_1) & 0 \\ 0 & A(u_1) & 0 & B(u_1) \\ C(u_1) & 0 & D(u_1) & 0 \\ 0 & C(u_1) & 0 & D(u_1) \end{pmatrix}, \quad \begin{pmatrix} A(u_2) & B(u_2) & 0 & 0 \\ C(u_2) & D(u_2) & 0 & 0 \\ 0 & 0 & A(u_2) & B(u_2) \\ 0 & 0 & C(u_2) & D(u_2) \end{pmatrix}$$

Recall that the $\mathcal{R}$-matrix is given by (78), which can be rewritten as

$$\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1 - u_2) = \begin{pmatrix} a(u_1 - u_2) & 0 & 0 & 0 \\ 0 & b(u_1 - u_2) & c(u_1 - u_2) & 0 \\ 0 & c(u_1 - u_2) & b(u_1 - u_2) & 0) \\ 0 & 0 & 0 & a(u_1 - u_2) \end{pmatrix} \text{ where } \begin{array}{l} a(\lambda) = \lambda + i \\ b(\lambda) = \lambda \\ c(\lambda) = i \end{array},$$

Then the fundamental commutation relation (87) establishes a relationship between the monodromy $\boldsymbol{\mathcal{T}}$-matrices and the $\mathcal{R}$-matrix. Then, plugging in the previous results yields

$$\begin{pmatrix} a_{12} & 0 & 0 & 0 \\ 0 & b_{12} & c_{12} & 0 \\ 0 & c_{12} & b_{12} & 0 \\ 0 & 0 & 0 & a_{12} \end{pmatrix} \begin{pmatrix} A(u_1)A(u_2) & A(u_1)B(u_2) & B(u_1)A(u_2) & B(u_1)B(u_2) \\ A(u_1)C(u_2) & A(u_1)D(u_2) & B(u_1)C(u_2) & B(u_1)D(u_2) \\ C(u_1)A(u_2) & C(u_1)B(u_2) & D(u_1)A(u_2) & D(u_1)B(u_2) \\ C(u_1)C(u_2) & C(u_1)D(u_2) & D(u_1)C(u_2) & D(u_1)D(u_2) \end{pmatrix}$$

$$= \begin{pmatrix} A(u_2)A(u_1) & B(u_2)A(u_1) & A(u_2)B(u_1) & B(u_2)B(u_1) \\ C(u_2)A(u_1) & D(u_2)A(u_1) & C(u_2)B(u_1) & D(u_2)B(u_1) \\ A(u_2)C(u_1) & B(u_2)C(u_1) & A(u_2)D(u_1) & B(u_2)D(u_1) \\ C(u_2)C(u_1) & D(u_2)C(u_1) & C(u_2)D(u_1) & D(u_2)D(u_1) \end{pmatrix} \begin{pmatrix} a_{12} & 0 & 0 & 0 \\ 0 & b_{12} & c_{12} & 0 \\ 0 & c_{12} & b_{12} & 0 \\ 0 & 0 & 0 & a_{12} \end{pmatrix}$$

$$\Rightarrow \Rightarrow \begin{pmatrix} a_{12}A_1A_2 & a_{12}A_1B_2 & a_{12}B_1A_2 & a_{12}B_1B_2 \\ b_{12}A_1C_2 + c_{12}C_1A_2 & b_{12}A_1D_2 + c_{12}C_1B_2 & b_{12}B_1C_2 + c_{12}D_1A_2 & b_{12}B_1D_2 + c_{12}D_1B_2 \\ c_{12}A_1C_2 + b_{12}C_1A_2 & b_{12}A_1D_2 + c_{12}C_1B_2 & c_{12}B_1C_2 + b_{12}D_1A_2 & c_{12}B_1D_2 + b_{12}D_1B_2 \\ a_{12}C_1C_2 & a_{12}C_1D_2 & a_{12}D_1C_2 & a_{12}D_1D_2 \end{pmatrix}$$

$$= \begin{pmatrix} a_{12}A_2A_1 & b_{12}B_2A_1 + c_{12}A_2B_1 & c_{12}B_2A_1 + b_{12}A_2B_1 & a_{12}B_2B_1 \\ a_{12}C_2A_1 & b_{12}D_2A_1 + c_{12}C_2B_1 & c_{12}D_2A_1 + b_{12}C_2B_1 & a_{12}D_2B_1 \\ a_{12}A_2C_1 & b_{12}B_2C_1 + c_{12}A_2D_1 & c_{12}B_2C_1 + b_{12}A_2D_1 & a_{12}B_2D_1 \\ a_{12}C_2C_1 & b_{12}D_2C_1 + c_{12}C_2D_1 & c_{12}D_2C_1 + b_{12}C_2D_1 & a_{12}D_2D_1 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} a_{12}[A_1, A_2] & a_{12}A_1B_2 - (b_{12}B_2A_1 + c_{12}A_2B_1) & a_{12}B_1A_2 - (c_{12}B_2A_1 + c_{12}A_2B_1) & a_{12}[B_1, B_2] \\ (b_{12}A_1C_2 + c_{12}C_1A_2) - a_{12}C_2A_1 & b_{12}[A_1, D_2] + c_{12}[C_1B_2] & (b_{12}[B_1, C_2] + c_{12}[D_1, A_2] & b_{12}B_1D_2 + c_{12}D_1B_2 - a_{12}D_2B_1 \\ c_{12}A_1C_2 + b_{12}C_1A_2 - a_{12}A_2C_1 & (b_{12}A_1D_2 + c_{12}C_1B_2) - b_{12}B_2C_1 + c_{12}A_2D_1) & c_{12}[B_1, C_2] + b_{12}[D_1, A_2] & c_{12}B_1D_2 + b_{12}D_1B_2 - a_{12}B_2D_1 \\ a_{12}[C_1, C_2] & a_{12}C_1D_2 - (b_{12}D_2C_1 + c_{12}C_2D_1) & a_{12}D_1C_2 - (c_{12}D_2C_1 + b_{12}C_2D_1) & a_{12}[D_1, D_2] \end{pmatrix} = \boldsymbol{0},$$

from which, it's been proven that $[\boldsymbol{\mathcal{T}}_{ij}(u_1), \boldsymbol{\mathcal{T}}_{ij}(u_2)] = 0, \ i, j = 1, 2$ as long as $a(u_1 - u_2) \neq 0$. From the other matrix-entries, it's immediately seen that

$$a(u_1 - u_2)B(u_1)A(u_2) = c(u_1 - u_2)B(u_2)A(u_1) + b(u_1 - u_2)A(u_2)B(u_1).$$

Swapping $u_1 \to u_2$ implies $u_1 - u_2 \to -(u_1 - u_2)$. Now, note that $c(u_1)/b(u_1) = g(u_1) = g(-u_1)$ and similarly $a(-u_1)/b(-u_1) = h(-u_1) \equiv f(u_1)$, which yields the second line of theorem 1. The third claim can then be similarly derived.

$$\textit{quod erat dem} \blacksquare$$

Eigenvalues of the Bethe Ansatz. Theorem 1's relations shows the model's inner structure, identifying an underlying algebraic structure somewhat similar to that of the Harmonic oscillator, where $A + D$ are associated with the eigenenergies, and where B and C can be interpreted as ladder operators.

Then, a Fock-like state space can be constructed for the N-body system, with its ground state satisfying $C\Omega^{\otimes N} = 0$. Then, $\omega_+^i = \begin{pmatrix} 1 & 0 \end{pmatrix}_i^\dagger$ describes i-th spin's state, with up spin. Then,

$$\mathbf{S}^+ \omega_+^i = 0,$$

Therefore, consider the $\mathcal{L}$-matrix, defined by (76), its action on a tensor product of states can be written out as

$$(97) \qquad \mathcal{L}_{N,a} v \otimes \omega_+^i = \begin{pmatrix} \lambda + \frac{i}{2} & \cdots \\ 0 & \lambda - \frac{i}{2} \end{pmatrix} v \otimes \omega_+^i, \ \forall v \in \mathbb{V}_a,$$

since $\mathbf{S}^+$ annihilates ω_+^i. Thus, the ground sate can be written as the physical ferromagnetic vacuum state with all its spins up. Said state is an eigenstate of $A(u)$ and $D(u)$ and is annihilated by $C(u)$, ie.

$$(98) \qquad \Omega^{\otimes N} = \bigotimes_{i=1}^{N} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \text{ where } \begin{aligned} A(u)\Omega^{\otimes N} &= \left(u + \tfrac{i}{2}\right)^N \Omega^{\otimes N}, \\ & \qquad\qquad\qquad\qquad C(u)\Omega^{\otimes N} = 0. \\ D(u)\Omega^{\otimes N} &= \left(u - \tfrac{i}{2}\right)^N \Omega^{\otimes N}, \end{aligned}$$

The $B(u)$ operators can then be used to construct the Bethe state

$$(99) \qquad |u_1, \cdots u_M\rangle = \prod_{j=1}^{M<N} B(u_j)\Omega^{\otimes N}, \text{ for some set } \{u_j\}_{j=1}^{M<N} \subset \mathbb{C}.$$

A priori, not all complex-valued sequences are allowed since the Bethe states must be eigenvectors of $A(u) + D(u)$, which yields some algebraic conditions.

Conditions on $\{u_j\}_{j=1}^{M<N}$. Consider the Bethe state given by equation (99). Then,

$$A(v)|u_1, \cdots u_M\rangle = A(v) \prod_{j=1}^{M<N} B(u_j)\Omega^{\otimes N}.$$

Using the fundamental commutation relations given theorem 1,

$$A(v)B(u) = f(v - u)B(u)A(v) + g(v - u)B(v)A(u).$$

Therefore,

$$(100) \qquad \begin{aligned} A(v)|u_1, \cdots u_M\rangle &= \left[f(v - u_1)B(u_1)A(v) + g(v - u_1)B(v)A(u_1)\right] \prod_{j=2}^{M<N} B(u_j)\Omega^{\otimes N} \\ &= \left[\prod_{k=1}^{\ell} f(v - u_k)\right] \alpha^N(v)|u_1, \cdots u_M\rangle + \mathcal{O}(A \times D), \end{aligned}$$

wherein, the first term is an scalar times the Bethe state, said scalar being the eigenvalue if and only if the second term, the term with products of A and D, cancels out. In that case, the Bethe state is an eigenvector of $A + D$. Then, continuing the previous calculation

(101)

$$A(v)\left|u_1, \cdots u_M\right\rangle = \left[f(v-u_1)B(u_1)A(v) + g(v-u_1)B(v)A(u_1)\right]B(u_2)\prod_{j=3}^{M<N}B(u_j)\Omega^{\otimes N}$$

$$= \left[f(v-u_1)f(v-u_2)B(u_1)B(u_2)A(v) + f(v-u_1)g(v-u_2)B(u_1)B(v)A(u_2)\right.$$

$$\left. + g(v-u_1)f(u_1-u_2)B(v)B(u_2)A(u_1) + g(v-u_1)g(v-u_2)B(v)B(u_2)\right]\prod_{j=3}^{M<N}B(u_j)\Omega^{\otimes N}.$$

Now, note that

$$\left(A(v)+D(v)\right)\left|u_1, \cdots u_M\right\rangle = \left[\left(\prod_{k=1}^{\ell}f(v-u_k)\right)\alpha^N(v) + \left(\prod_{k=1}^{\ell}h(v-u_k)\right)\delta^N(v)\right]\left|u_1, \cdots u_M\right\rangle$$

$$+ \sum_{k=1}^{\ell}\left[\left(g(v-u_k)\left[\prod_{j\neq k}^{\ell}f(u_k-u_j)\right]\alpha^N(u_k)\right) + h(v-u_k)\left(\prod_{j\neq k}^{\ell}h(u_k-u_j)\right)\delta^N(u_j)\right.$$

$$\left. \times \prod_{\substack{j=1 \\ j\neq k}}^{\ell}\Omega^{\otimes N}B(u_i)\right]$$

$$\equiv \left[\left(\prod_{k=1}^{\ell}f(v-u_k)\right)\alpha^N(v) + \left(\prod_{k=1}^{\ell}h(v-u_k)\right)\delta^N(v)\right]\left|u_1, \cdots u_M\right\rangle + 0$$

$$\Rightarrow \sum_{k=1}^{\ell}\left[\left(g(v-u_k)\left[\prod_{j\neq k}^{\ell}f(u_k-u_j)\right]\alpha^N(u_k)\right) + h(v-u_k)\left(\prod_{j\neq k}^{\ell}h(u_k-u_j)\right)\delta^N(u_j)\right] = 0$$

$$\sum_{k=1}^{\ell}\left[\left(g(v-u_k)\left[\prod_{j\neq k}^{\ell}f(u_k-u_j)\right]\alpha^N(u_k)\right) - g(v-u_k)\left(\prod_{j\neq k}^{\ell}h(u_k-u_j)\right)\delta^N(u_j)\right] = 0$$

$$\sum_{k=1}^{\ell}g(v-u_k)\left[\prod_{j\neq k}^{\ell}f(u_k-u_j)\right]\alpha^N(u_k) + \prod_{j\neq k}^{\ell}h(u_k-u_j)\delta^N(u_j)\right] = 0,$$

which must be true for all k,

(102)
$$\left[\prod_{j\neq k}^{\ell}f(v-u_k)\right]\alpha^N(u_j) = \left[\prod_{j\neq k}^{\ell}h(v-u_k)\right]\delta^N(u_j).$$

Now using theorem 1 functions and the that $\alpha(u) = u + \frac{i}{2}$ and $\delta(u) = u - \frac{i}{2}$, the previous equation can be rewritten as

$$\left[\prod_{j\neq k}^{\ell}f(u_k-u_j)\right]\alpha^N(u_j) = \left[\prod_{j\neq k}^{\ell}h(u_k-u_j)\right]\delta^N(u_j)$$

(103)
$$\prod_{j\neq k}^{\ell}\frac{u_k-u_j-i}{u_k-u_j}\left(u_j+\frac{i}{2}\right) = \prod_{j\neq k}^{\ell}\frac{u_k-u_j+i}{u_k-u_j}\left(u_j-\frac{i}{2}\right)$$

$$0 = \prod_{j\neq k}^{\ell}\frac{u_k-u_j-i}{u_k-u_j}\left(u_j+\frac{i}{2}\right) - \prod_{j\neq k}^{\ell}\frac{u_k-u_j+i}{u_k-u_j}\left(u_j-\frac{i}{2}\right),$$

which can be rearranged to yield the Bethe ansatz equations,

(104)
$$\left(\frac{u_k + \frac{i}{2}}{u_k - \frac{i}{2}}\right)^N = \prod_{j \neq k}^{\ell} \frac{u_k - u_j + i}{u_k - u_j - i},$$

which is a set of ℓ non linear equations on the coefficients $u_k \in \mathbb{C}$.

Momentum and Energy. Let $\{u_k\}_{k=1}^{\ell < N} \subset \mathbb{C}$ satisfy the Bethe equations. Then,

(105)

$$(A(v) + D(v)) |u_1, \cdots u_M\rangle = \left[\left(\prod_{k=1}^{\ell} f(v - u_k)\right) \alpha^N(v) + \left(\prod_{k=1}^{\ell} h(v - u_k)\right) \delta^N(v)\right] |u_1, \cdots u_M\rangle$$

$$(A(v) + D(v)) |u_1, \cdots u_M\rangle = \left[\left(v + \frac{i}{2}\right)^N \prod_{k=1}^{\ell} \frac{v - u_k - i}{v - u_k} + \left(v - \frac{i}{2}\right)^N \prod_{k=1}^{\ell} \frac{v - u_k + i}{v - u_k}\right] |u_1, \cdots u_M\rangle$$

$$\left(A\left(\frac{i}{2}\right) + D\left(\frac{i}{2}\right)\right) |u_1, \cdots u_M\rangle = i^N \prod_{k=1}^{\ell} \frac{u_k + \frac{i}{2}}{u_k - \frac{i}{2}} |u_1, \cdots u_M\rangle \equiv \lambda(\{u_k\}_{k=1}^{\ell < N}) |u_1, \cdots u_M\rangle$$

Therefore, via the transfer matrix, it holds that $U = e^{iP} = i^{-N} \mathfrak{t}\left(\frac{i}{2}\right) = i^{-N}\left(A\left(\frac{i}{2}\right) + D\left(\frac{i}{2}\right)\right)$, from which

(106)
$$P |u_1, \cdots u_\ell\rangle = \sum_{k=1}^{\ell} p(u_k) |u_1, \cdots u_\ell\rangle \text{ where } p(u_k) = \frac{1}{i} \log \frac{u_k + \frac{i}{2}}{u_k - \frac{i}{2}}.$$

Now, equation (95) establishes that the XXX Hamiltonian belongs to the monodromy matrices' family. Furthermore, it provides a link between the $A + D$'s eigenvalues and the eigenenergies. In effect, said eigenenergies may be found by differentiating the $A + D$'s eigenvalues over u and setting $u = z_0 = \frac{i}{2}$. Let the former

$$\Lambda(v) = \left(v + \frac{i}{2}\right)^N \prod_{k=1}^{\ell} \frac{v - u_k - i}{v - u_k} + \left(v - \frac{i}{2}\right)^N \prod_{k=1}^{\ell} \frac{v - u_k + i}{v - u_k},$$

then,

$$\frac{d \log \Lambda(v)}{dv}\bigg|_{v=z_0} = \frac{1}{\Lambda(v)} \frac{d\Lambda(v)}{dv}\bigg|_{v=z_0}$$

$$= \frac{1}{\Lambda(v)\big|_{v=z_0}} \left[N i^{N-1} \prod_{k=1}^{\ell} \frac{-u_k - \frac{i}{2}}{\frac{i}{2} - u_k} + i^N \sum_k \frac{i}{(\frac{i}{2} - u_k)^2} \prod_{\substack{j=1 \\ j \neq k}}^{\ell} \frac{-u_j - \frac{i}{2}}{\frac{i}{2} - u}\right]$$

$$= i \left(\sum_{k=1}^{\ell} \frac{1}{u_k^2 + \frac{1}{4}} - N\right).$$

Then, using equation (95), yields the eigenenergies

(107)
$$H |u_1, \cdots u_M\rangle = \left[\frac{i}{2} i \left(\sum_{k=1}^{\ell} \frac{1}{u_k^2 + \frac{1}{4}} - N\right) - \frac{N}{2}\right] |u_1, \cdots u_M\rangle$$

or, equivalently,

(108)
$$\mathbf{H} |u_1, \cdots u_M\rangle = \sum_{k=1}^{\ell} \epsilon(u_k) |u_1, \cdots u_M\rangle \text{ where } \epsilon(u_k) = -\frac{1}{2} \frac{1}{u_k^2 + \frac{1}{4}},$$

which concludes the Bethe solution for the XXX Heisenberg chain.

2.5. **Numerical solution to Fermionic models.** Consider a Hamiltonian describing a fermionic system, given by

$$(109) \qquad \mathbf{H} = J \sum_{j=1}^{L-1} (f_j^\dagger f_{j+1} + f_{j+1}^\dagger f_j) + \sum_{j=1} \lambda_j f_j^\dagger f_j, \quad \begin{matrix} \text{with the usual} \\ \text{conmutation rules} \end{matrix} \quad \begin{matrix} \{f_j, f_k\} = \{f_j^\dagger, f_k^\dagger\} = 0 \\ \{f_j, f_k^\dagger\} = \delta_{jk} \end{matrix}$$

where L indicates the number of lattice sites, J is the hopping strength, which could be either positive or negative, and where λ_j is the on-site potential strength[13]. Said Hamiltonian has open boundaries conditions since there is no hopping term across the boundary. Note that we can rewrite (109) as

$$(110) \qquad \mathbf{H} = \sum_{i,j=1}^L \mathcal{M}_{ij} f_i^\dagger f_j \text{ with } \mathcal{M}_{ij} = \begin{cases} \mathcal{M} \in \mathrm{GL}(L,\mathbb{R}), \\ \lambda_i & \text{if } i = j \\ J & \text{if } j = i+1 \text{ or } i = j+1 \\ 0 & \text{otherwise} \end{cases},$$

which is a positive-defined tri-diagonal matrix. Let $\mathbf{f} = (f_1 \ f_2 \ \cdots f_L)^{\mathrm{T}}$ be a vector of the L fermionic operators. Then, (109) can be rewritten as

$$(111) \qquad \mathbf{H} = \mathbf{f}^\dagger \mathcal{M} \mathbf{f}.$$

Since $\mathcal{M}$ is symmetric, then it can be diagonalized $\mathcal{M} = A\mathcal{D}A^{\mathrm{T}}$, where $A \in \mathbb{R}^{L\times L}$ is a real orthogonal matrix and with $\mathcal{D}_{ij} \in \mathbb{R}^{L\times L} \mid \mathcal{D}_{ij} = \epsilon_i \delta_{ij}$. In this context, the A-matrix acts on the fermionic operator as a Bogoliubov transformation, allowing for (109) to be rewritten as

$$(112) \qquad \mathbf{H} = \mathbf{f}^\dagger A\mathcal{D}A^{\mathrm{T}}\mathbf{f} = \mathbf{d}^\dagger \mathcal{D} \mathbf{d}$$

where $\mathbf{d} = A^{\mathrm{T}}\mathbf{f}$. Since the A-matrix is orthogonal, the new d_k-operators are fermionic operators as well, satisfying (109)'s anti-commutation rules. Then, the new fermionic operators are

$$\begin{matrix} d_k = \sum_{j=1}^L A_{jk} f_j \\ f_j = \sum_{k=1}^L A_{jk} d_k \end{matrix} \quad \text{since } A^{\mathrm{T}}A = \sum_{j,k=1}^L A_{jk} A_{kj} = \mathbb{1}_L.$$

Then, we can expand (112) in terms of the lattices, as follows

$$(113) \qquad \mathbf{H} = \sum_{k=1}^L \epsilon_k d_k^\dagger d_k,$$

which is a sum of number operators with potentials. The eigenstates can then be constructed from the the theory's vacuum state, by applying the $d_k^\dagger$-fermionic operators. In the Heisenberg-picture, the d_k-operators can be evolved via the Heisenberg equation of motion

$$(114) \qquad \frac{d}{dt} d_k = i[\mathbf{H}, d_k],$$

and using that $d_k^2 = 0$, it turns out that (114)'s solution is simply $d_k(t) = e^{-i\epsilon_k t} d_k$. The system's correlation can be easily found by analyzing the following matrix. Let $\mathcal{N}_{jk} = \langle d_j^\dagger d_k \rangle$, where the expectation value is taken via calculating the operator's trace along the Fock space, which takes the following values

---

[13]The λ_j-term frequently appears in many condensed matter models, with different numerical values and interpretations, eg.

- In the XX model, $\lambda_j = \lambda \ \forall j$.
- While for the Anderson model $\lambda_j \in \mathcal{U}_{\mathbb{R}_{[-W,W]}}$, a uniform random variable, with W being the disorder strength.
- In the Aubry-André model, $\lambda_j = \lambda \cos(2\pi\sigma j)$, with $\sigma \in \mathbb{I}$ and λ quantifying the disorder strength.

$$(115) \qquad \mathcal{N}_{jk} = \langle d_j^\dagger d_k \rangle = \begin{cases} 0 \text{ or } 1 \text{ if } j = k \\ \quad 0 \text{ if } j \neq k \end{cases},$$

ie. different lattice-sites are not correlated and there can only be a single fermion at most per lattice site, in accordance with Pauli's principle. A ground state, for example, would choose to turn on all fermions in the eigenmode d-space such that $\epsilon_k < 0$. If instead, the expectation value is taken with thermal states, the Fermi-Dirac distribution is returned,

$$(116) \qquad \mathcal{N}_{jk} = \langle d_j^\dagger d_k \rangle_{\text{th}} = \frac{1}{1 + e^{\beta \epsilon_k + \mu}} \delta_{jk}.$$

Another interesting quality is a system with an initial configuration where the system's initial state, in real space, is known. In this setting, $\mathcal{N}_{jk}$ is known for all lattices. Consider for example the Anderson model, where the system's initial state is given by a single tensor product of n-fermionic states in real space, with $n < L$. Then, for all lattice sites, we have that $\mathcal{N}_{jj}$ is either zero or one. The $\mathcal{N}_{jk}$-matrix entries can then be evaluated as

$$\langle d_j^\dagger d_k \rangle = \sum_{i,j=1}^{n<L} A_{ik} A_{jl} \langle f_i^\dagger f_j \rangle = \sum_{j=1} A_{jk} A_{jl} \langle f_j^\dagger f_j \rangle,$$

which can then be numerically computed to obtain the LHS expectation value. In general, this $\mathcal{N}_{jk}$-matrix will not be diagonal, which is reasonable since the system's real configuration is not an eigenstate. In principle and in practice, by inverting (117), we can evolve any number operator or two-body correlation operator, ie.

$$(117) \qquad \langle f_j^\dagger f_k \rangle = \sum_{k,l=1}^{n} A_{jk} A_{jl} \langle d_k^\dagger d_l \rangle.$$

This quantities' time evolution can then be found out to be

$$(118) \qquad \langle f_j^\dagger(t) f_k(t) \rangle = \sum_{k,l=1}^{L} e^{i(\epsilon_k - \epsilon_l)t} A_{jk} A_{jl} \langle d_k^\dagger d_l \rangle,$$

which can then be numerically solved.

3. Quantum Field-Therotic approach to Condensed Matter Physics

3.1. **Mean field theory. From "Many Body Quantum Theory in Condensed Matter Physics" by Henrik Bruus and Karsten Flensberg.**

The physics of interacting particles is often very complicated because the motion of the individual particles depend on the position of all other, thus these variables become correlated. This is clearly the case for a system of charged particles interacting by Coulomb forces, such as eg. the electron gas. There, the probability to find two electrons in close proximity is expected to be small due to the strong repulsive interaction. Consequently, due to these correlation effects, there is a suppressed density in the neighbourhood of every electron, and there is a correlation hole. Nevertheless, in spite of these complications, there a re a number of cases where a more crude treatment, not fully including the correlations, gives a good physical model. In these cases, it suffices to include correlations o the average. This means that the effect of the other particles is included as a mean density (or mean field), leaving an effective single particle problem, which is solvable. These mean fields are chosen so that they minimize the free energy, which in turn, ensures that the method is self-consistent. This method is called mean field theory, which allows to neglect detailed dynamics and time-dependent second quantization methods.

First, consider a system with two types of particles, described by operators $\mathbf{a}_\nu, \mathbf{b}_\mu$, respectively. In this model, the only interactions present are those between different kind of particles. The Hamiltonian then reads

$$\mathbf{H}_0 = \sum_\nu \varepsilon_\nu^a \mathbf{a}_\nu^\dagger \mathbf{a}_\nu + \sum_\nu \varepsilon_\mu^b \mathbf{b}_\mu^\dagger \mathbf{b}_\mu,$$

(119) $\qquad \mathbf{H} = \mathbf{H}_0 + \mathbf{V}_{\text{int}}, \text{ with}$

$$\mathbf{V}_{\text{int}} = \sum_{\nu\nu',\mu\mu'} \mathcal{V}_{\nu\mu,\nu'\mu'} \mathbf{a}_\nu^\dagger \mathbf{b}_\mu^\dagger \mathbf{b}_{\mu'} \mathbf{a}_{\nu'}.$$

Now it may be expected that the density operator $\mathbf{a}_\nu^\dagger \mathbf{a}_{\nu'}$ and $\mathbf{b}_\mu^\dagger \mathbf{b}_{\mu'}$, deviate only little from their averages values, $\langle \mathbf{a}_\nu^\dagger \mathbf{a}_{\nu'} \rangle$ and $\langle \mathbf{b}_\nu^\dagger \mathbf{b}_{\nu'} \rangle$, respectively. It is natural then to use this deviation as a small parameter and perform an expansion. Let the deviation operators be

(120) $\qquad \mathbf{d}_{\nu\nu'} = \mathbf{a}_\nu^\dagger \mathbf{a}_{\nu'} - \langle \mathbf{a}_\nu^\dagger \mathbf{a}_{\nu'} \rangle, \qquad\qquad \mathbf{e}_{\mu\mu'} = \mathbf{b}_\nu^\dagger \mathbf{b}_{\nu'} - \langle \mathbf{b}_\mu^\dagger \mathbf{b}_{\mu'} \rangle.$

Then, in terms of these new operators, the Hamiltonian reads

$$\mathbf{H}_{\text{MFT}} = \mathbf{H}_0 + \mathbf{V}_{\text{MFT}} + \underbrace{\sum_{\nu\nu',\mu\mu'} \mathbf{d}_{\nu\nu'} \mathbf{e}_{\mu\mu'}}_{}{}^{\,0},$$

where

(121) $\qquad \mathbf{V}_{\text{MFT}} = \sum_{\nu\nu',\mu\mu'} \mathcal{V}_{\nu\mu,\nu'\mu'} \left[\mathbf{a}_\nu^\dagger \mathbf{a}_{\nu'} \langle \mathbf{b}_\mu^\dagger \mathbf{b}_{\mu'} \rangle + \mathbf{b}_\mu^\dagger \mathbf{b}_{\mu'} \langle \mathbf{a}_\nu^\dagger \mathbf{a}_{\nu'} \rangle - \langle \mathbf{a}_\nu^\dagger \mathbf{a}_{\nu'} \rangle \langle \mathbf{b}_\mu^\dagger \mathbf{b}_{\mu'} \rangle \right] + \mathcal{O}(de).$

The mean field Hamiltonian H_{MFT} thus contains only single-particle operators, and thus the original many-body problem has been reduced to a single-particle problem, which can always be solved[14]

Mean field theory may be formulated in a different way: if an interaction term involving the product of two operators, $\mathbf{A}, \mathbf{B}$, ie. $\mathbf{H}_{AB} = \mathbf{AB}$, then the mean field approximation is given by

$$\mathbf{H}_{AB}^{\text{MFT}} = \mathbf{A}\langle \mathbf{B} \rangle + \langle \mathbf{A} \rangle \mathbf{B} - \langle \mathbf{A} \rangle \langle \mathbf{B} \rangle.$$

---

[14]In effect, any quadratic Hamiltonian may be diagonalized since it can be written as $\mathbf{H} = \sum_{ij} a_i^\dagger \mathcal{H}_{ij} a_i^\dagger$. Since this $\mathcal{H}$ is hermitian, then there exists a transformation

$$\alpha_i \to \sum_j U_{ij} a_j, \; U \in U(N),$$

such that $\mathcal{H}$ is diagonal: $\sum_{kk'} U_{ik}^\dagger h_{kk'} U_{k'j} = \delta_{ij} \varepsilon_i$. Then, in terms of this new basis, the Hamiltonian then becomes $\mathbf{H} = \sum_i \alpha_i^\dagger \varepsilon_i \alpha_i$, and $\{\varepsilon_i\}$ are thus $\mathcal{H}$'s eigenvalues.

The question is then, how to find the averages $\langle \mathbf{a}_\nu^\dagger \mathbf{a}_{\nu'} \rangle$ and $\langle \mathbf{b}_\mu^\dagger \mathbf{b}_{\mu'} \rangle$. In principle, there are two possible routes.

- Method 1: these averages are to be determined self-consistently, ie. when calculating the averages

(122)
$$\bar{\mathbf{n}}_{\nu\nu'}^a = \langle \mathbf{a}_\nu^\dagger \mathbf{a}_{\nu'} \rangle, \qquad\qquad \bar{\mathbf{n}}_{\mu\mu'}^b = \langle \mathbf{b}_\mu^\dagger \mathbf{b}_{\nu'} \rangle,$$

using the new mean field Hamiltonian, the same answer should be re-obtained. This entails the following prescriptions for $\bar{\mathbf{n}}_{\nu\nu'}^a$ and $\bar{\mathbf{n}}_{\mu\mu'}^b$,

(123)
$$\begin{aligned} \bar{\mathbf{n}}_{\nu\nu'}^a &= \langle \mathbf{a}_\nu^\dagger \mathbf{a}_{\nu'} \rangle_{\text{MFT}} = \frac{1}{\mathcal{Z}_{\text{MFT}}} \operatorname{Tr}\left(e^{-\beta \mathbf{H}_{\text{MFT}}} \mathbf{a}_\nu^\dagger \mathbf{a}_{\nu'} \right) \\ \bar{\mathbf{n}}_{\nu\nu'}^b &= \langle \mathbf{b}_\mu^\dagger \mathbf{b}_{\mu'} \rangle_{\text{MFT}} = \frac{1}{\mathcal{Z}_{\text{MFT}}} \operatorname{Tr}\left(e^{-\beta \mathbf{H}_{\text{MFT}}} \mathbf{b}_\mu^\dagger \mathbf{b}_{\mu'} \right), \end{aligned} \quad \text{where } \mathcal{Z}_{\text{MFT}} = \operatorname{Tr}\left(e^{-\beta \mathbf{H}_{\text{MFT}}} \right).$$

Said equations are the self-consistency equations, since the averages are given in terms of the mean field hamiltonian and the mean field partition function, which themselves depend on these averages.

- Method 2: the alternative route consists on using the $\mathbf{n}_{\nu\nu'}$ which minimizes the free energy F_{MFT} of the mean field Hamiltonian. Then,

(124)
$$\begin{aligned} 0 &= \frac{d}{d\bar{\mathbf{n}}_{\nu\nu'}^a} F_{\text{MFT}} = \frac{d}{d\bar{\mathbf{n}}_{\nu\nu'}^a}\left(-\frac{\log \mathcal{Z}_{\text{MFT}}}{\beta} \right) = \frac{1}{\mathcal{Z}_{\text{MFT}}} \operatorname{Tr}\left(e^{-\beta \mathbf{H}_{\text{MFT}}} \frac{d}{d\bar{\mathbf{n}}_{\nu\nu'}^a} \mathbf{H}_{\text{MFT}} \right) \\ &= \frac{1}{\mathcal{Z}_{\text{MFT}}} \operatorname{Tr}\left(e^{-\beta \mathbf{H}_{\text{MFT}}} \left(\sum_{\nu\nu',\mu\mu'} \mathcal{V}_{\nu\mu,\nu'\mu'} (\mathbf{b}_\mu^\dagger \mathbf{b}_\mu - \bar{\mathbf{n}}_{\mu\mu'}^b) \right) \right) \\ &= \frac{1}{\mathcal{Z}_{\text{MFT}}} \operatorname{Tr}\left(e^{-\beta \mathbf{H}_{\text{MFT}}} \sum_{\nu\nu',\mu\mu'} \mathcal{V}_{\nu\mu,\nu'\mu'} \mathbf{b}_\mu^\dagger \mathbf{b}_\mu \right) - \frac{1}{\mathcal{Z}_{\text{MFT}}} \operatorname{Tr}\left(e^{-\beta \mathbf{H}_{\text{MFT}}} \sum_{\nu\nu',\mu\mu'} \mathcal{V}_{\nu\mu,\nu'\mu'} \bar{\mathbf{n}}_{\mu\mu'}^b \right) \\ &= \sum_{\mu\mu'} \frac{1}{\mathcal{Z}_{\text{MFT}}} \mathcal{V}_{\nu\mu,\nu'\mu'} \operatorname{Tr}\left(e^{-\beta \mathbf{H}_{\text{MFT}}} \mathbf{b}_\mu^\dagger \mathbf{b}_\mu \right) - \sum_{\mu\mu'} \mathcal{V}_{\nu\mu,\nu'\mu'} \bar{\mathbf{n}}_{\mu\mu'}^b \frac{1}{\mathcal{Z}_{\text{MFT}}} \operatorname{Tr}\left(e^{-\beta \mathbf{H}_{\text{MFT}}} \right) \\ &= \sum_{\mu\mu'} \mathcal{V}_{\nu\mu,\nu'\mu'} \left(\langle \mathbf{b}_\mu^\dagger \mathbf{b}_{\mu'} \rangle_{\text{MFT}} - \bar{\mathbf{n}}_{\mu\mu'}^b \right), \end{aligned}$$

which should hold for any pair (ν, ν') and hence, the last parenthesis should vanish. Thus, the self-consistency equation is obtained as well for the second method, making both approaches equivalent.

Consider the average interaction energy, $\langle \mathbf{V}_{\text{int}} \rangle$, a natural approximation would be to evaluate the expectation value of the $\mathbf{a}$ and $\mathbf{b}$ operators separately, this is

$$\langle \mathbf{V}_{\text{int}} \rangle \approx \sum_{\nu\nu',\mu\mu'} \mathcal{V}_{\nu\mu,\nu'\mu'} \langle \mathbf{a}_\nu^\dagger \mathbf{a}_{\nu'} \rangle \langle \mathbf{b}_\mu^\dagger \mathbf{b}_{\mu'} \rangle,$$

which is equivalent to assuming that the $\mathbf{a}$ and $\mathbf{b}$ operators are uncorrelated. This is, in essence, the approximation done in the mean-field approach. In effect, the interaction term can be evaluated using the mean field Hamiltonian

$$\langle \mathbf{V}_{\text{int}} \rangle_{\text{MFT}} = \frac{1}{\mathcal{Z}_{\text{MFT}}} \operatorname{Tr}\left((e^{-\beta \mathbf{H}_{\text{MFT}}} \mathbf{V}_{\text{int}}) \right).$$

Since, the mean field Hamiltonian can be separated into a part containing only $\mathbf{a}$-operators and another term containing only $\mathbf{b}$-operators, $\mathbf{H}_{\text{MFT}} = \mathbf{H}_{\text{MFT}}^a + \mathbf{H}_{\text{MFT}}^b$, the average factorization yields

(125)
$$\langle \mathbf{V}_{\text{int}} \rangle_{\text{MFT}} = \sum_{\mu\mu'} \mathcal{V}_{\nu\mu,\nu'\mu'} \langle \mathbf{a}_\nu^\dagger \mathbf{a}_{\nu'} \rangle_{\text{MFT}} \langle \mathbf{b}_\mu^\dagger \mathbf{b}_{\mu'} \rangle_{\text{MFT}}.$$

The mean field theory approach hence provides a consistent a physically sensible method to study interacting system where correlation are less important. Mean field theory is only valid provided d is actually small. This

can be checked by calculating $\langle \mathbf{d} \rangle$, using the neglected term in the Hamiltonian (that concerning the cross term of $\mathbf{d}_{\nu\nu'}\mathbf{e}_{\mu\mu'}$) as a perturbation, and then compare the result to $\langle \mathbf{a}_\nu^\dagger \mathbf{a}_{\nu'} \rangle$. If the difference is not negligible, then either the wrong mean field parameter has been chosen or the method simply fails and other more adequate tools must be employed.

In practice, it is necessary to assume something about the averages $\langle \mathbf{a}_\nu^\dagger \mathbf{a}_{\nu'} \rangle$ and $\langle \mathbf{b}_\mu^\dagger \mathbf{b}_{\mu'} \rangle$. Even though, there is a prescription for which to find which averages are important, there are simply too many combinations. Suppose there are N different quantum numbers, then there are in principle N^2 combinations, which then yields N^2 coupled non-linear equations, which is only tractable for small systems. Often, symmetry arguments can help in reducing the number of parameters. Suppose for example that the system's Hamiltonian has translational symmetry, such that the momentum space is a natural choice. For a system of particles described by $(\mathbf{c}\ \mathbf{c}^\dagger)$-operators, then

$$(126) \qquad \langle \mathbf{c}_\mathbf{k}^\dagger \mathbf{c}_{\mathbf{k}'} \rangle = \frac{1}{\mathcal{V}} \int_{\mathbb{R}^6} d\mathbf{x}d\mathbf{x}' e^{-i\mathbf{k}'\cdot\mathbf{x}'} e^{i\mathbf{k}\cdot\mathbf{x}} \langle \Psi^\dagger(\mathbf{x})\Psi(\mathbf{x}') \rangle.$$

It is natural to assume that the system is homogenous, which means that

$$(127) \qquad \langle \Psi^\dagger(\mathbf{x})\Psi(\mathbf{x}') \rangle = f(\mathbf{x} - \mathbf{x}') \Rightarrow \langle \mathbf{c}_\mathbf{k}^\dagger \mathbf{c}_{\mathbf{k}'} \rangle = \langle \mathbf{n}_\mathbf{k} \rangle \delta_{\mathbf{k}\mathbf{k}'}.$$

The homogeneity assumption is however not always true, since in some cases the system's symmetry is lower than that of the Hamiltonian. Eg. if the system spontaneously orders into a state with a spatial density variation, like a wave formation, then the average $\langle \Psi^\dagger(\mathbf{x})\Psi(\mathbf{x}') \rangle$ is not a function of $\mathbf{x} - \mathbf{x}'$ only. Instead it has a lower and more restricted symmetry, namely that

$$\langle \Psi^\dagger(\mathbf{x})\Psi(\mathbf{x}') \rangle = h(\mathbf{x}, \mathbf{x}') \text{ with } h(\mathbf{x}, \mathbf{x}') = h(\mathbf{x} + \mathbf{R}, \mathbf{x}' + \mathbf{R}), \ \mathbf{R} \text{ a lattice vector.}$$

This kind of crystal structure exists in nature and when it happens, the system undergoes a broken symmetry phenomena. It is important to realize that this solution cannot be found by taking the homogeneity assumption but instead by calculating $\langle \mathbf{c}_\mathbf{k}^\dagger \mathbf{c}_{\mathbf{k}+\mathbf{Q}} \rangle$, leading to the possibility of this quantity being finite, where $\mathbf{R} \cdot \mathbf{Q} = 2\pi$. Thus, the choice of the proper mean field parameters requires physical motivation about which possible states one expects.

Hartree-Fock theory. Mean field theory's methods, upto now, concern only interaction between different particles, wherein correlation terms are replaced by their average values plus small corrections. For interactions between identical particles this, however, does not exhaust the possibilities and only includes the so-called Hartree term, requiring a more general approximation scheme, given by the Hartree-Fock approximation. Consider a quantum system of interacting particles, whose Hamiltonian is given by

$$(128) \qquad \mathbf{H} = \mathbf{H}_0 + \mathbf{V}_{\text{int}}, \text{ with} \qquad \begin{aligned} \mathbf{H}_0 &= \sum_\nu \varepsilon_\nu \mathbf{c}_\nu^\dagger \mathbf{c}_\nu \\ \mathbf{V}_{\text{int}} &= \tfrac{1}{2} \sum_{\nu\nu',\mu\mu'} \mathcal{V}_{\nu\mu,\nu'\mu'} \mathbf{c}_\nu^\dagger \mathbf{c}_\mu^\dagger \mathbf{c}_{\mu'} \mathbf{c}_{\nu'}. \end{aligned}$$

The theory's main idea is that the operator $\rho_{\mu\mu'} = \mathbf{c}_\mu^\dagger \mathbf{c}_{\mu'}$ is large only when the average $\langle \rho_{\mu\mu'} \rangle$ is non-zero. For most $(\nu\nu')$-combinations, the average $\langle \rho_{\nu\nu'} \rangle$ is zero. Therefore, the four-body interaction term may be rewritten in terms of a deviation from the average value

$$\mathbf{c}_\nu^\dagger \left(\mathbf{c}_\mu^\dagger \mathbf{c}_{\mu'} - \langle \mathbf{c}_\mu^\dagger \mathbf{c}_{\mu'} \rangle \right) \mathbf{c}_{\nu'} + \mathbf{c}_\nu^\dagger \mathbf{c}_{\nu'} \langle \mathbf{c}_\mu^\dagger \mathbf{c}_{\mu'} \rangle.$$

If $\nu' \neq \mu$, then $\mathbf{c}_\nu$ operator commutes with the parenthesis, which is true except in a set of measure zero. With this assumption, then

(129)

$$\mathbf{c}_\nu^\dagger \left(\mathbf{c}_\mu^\dagger \mathbf{c}_{\mu'} - \langle \mathbf{c}_\mu^\dagger \mathbf{c}_{\mu'} \rangle \right) \mathbf{c}_{\nu'} + \mathbf{c}_\nu^\dagger \mathbf{c}_{\nu'} \langle \mathbf{c}_\mu^\dagger \mathbf{c}_{\mu'} \rangle \stackrel{\mu \neq \nu'}{=} \mathbf{c}_\nu^\dagger \mathbf{c}_{\nu'} \left(\mathbf{c}_\mu^\dagger \mathbf{c}_{\mu'} - \langle \mathbf{c}_\mu^\dagger \mathbf{c}_{\mu'} \rangle \right) + \mathbf{c}_\nu^\dagger \mathbf{c}_{\nu'} \langle \mathbf{c}_\mu^\dagger \mathbf{c}_{\mu'} \rangle$$
$$= \left(\mathbf{c}_\nu^\dagger \mathbf{c}_{\nu'} - \langle \mathbf{c}_\nu^\dagger \mathbf{c}_{\nu'} \rangle \right) \left(\mathbf{c}_\mu^\dagger \mathbf{c}_{\mu'} - \langle \mathbf{c}_\mu^\dagger \mathbf{c}_{\mu'} \rangle \right) + \mathbf{c}_\nu^\dagger \mathbf{c}_{\nu'} \langle \mathbf{c}_\mu^\dagger \mathbf{c}_{\mu'} \rangle - \mathbf{c}_\mu^\dagger \mathbf{c}_{\mu'} \langle \mathbf{c}_\nu^\dagger \mathbf{c}_{\nu'} \rangle - \langle \mathbf{c}_\mu^\dagger \mathbf{c}_{\mu'} \rangle \langle \mathbf{c}_\nu^\dagger \mathbf{c}_{\nu'} \rangle.$$

The first term is proportional to the deviation squared and then may be neglected, yielding the Hartree approximation for interactions

(130)
$$\mathbf{V}_{\text{int}}^{\text{HF}} = \frac{1}{2} \sum_{\nu\nu',\mu\mu'} \mathcal{V}_{\nu\mu,\nu'\mu'} \left[\bar{\mathbf{n}}_{\mu\mu'} \mathbf{c}_\nu^\dagger \mathbf{c}_{\nu'} + \bar{\mathbf{n}}_{\nu\nu'} \mathbf{c}_\mu^\dagger \mathbf{c}_{\mu'} - \bar{\mathbf{n}}_{\nu\nu'} \bar{\mathbf{n}}_{\mu\mu'} \right],$$

which is the same result that would be obtained if the $(\nu\nu')$ and $(\mu\mu')$-operators were considered to be from different kinds of particles. However, this is not yet the full result, since the μ and ν-labelled operators represent identical, indistinguishable particles and there is therefore one combination not accounted for by the Hartree approximation, the so-called Fock exchange term. This new term appears since the product of four one-body operators also gives a large contribution if $\langle \mathbf{c}_\nu^\dagger \mathbf{c}_{\nu'} \rangle$ finite. To derive the mean field contribution from this possibility the $\langle \mathbf{c}_\nu^\dagger \mathbf{c}_{\mu'} \rangle$ term must first be replaced by its average value and do the same with the combination of $\langle \mathbf{c}_\mu^\dagger \mathbf{c}_{\nu'} \rangle$, thus yielding

(131)
$$\mathbf{V}_{\text{int}}^{\text{Fock}} = -\frac{1}{2} \sum_{\nu\nu',\mu\mu'} \mathcal{V}_{\nu\mu,\nu'\mu'} \left[\bar{\mathbf{n}}_{\nu\mu'} \mathbf{c}_\mu^\dagger \mathbf{c}_{\nu'} + \bar{\mathbf{n}}_{\mu\nu'} \mathbf{c}_\nu^\dagger \mathbf{c}_{\mu'} - \bar{\mathbf{n}}_{\mu\nu'} \bar{\mathbf{n}}_{\nu\mu'} \right].$$

Therefore, the final mean field Hamiltonian within the Hartree-Fock approximation is

(132)
$$\mathbf{H}^{\text{HF}} = \mathbf{H}_0 + \mathbf{V}_{\text{int}}^{\text{Hartree}} + \mathbf{V}_{\text{int}}^{\text{Fock}}.$$

Homogeneous electron gas. Consider the example of a translationally invariant homogeneous electron gas, wherein the expectation value $\langle \mathbf{c}_\mathbf{k}^\dagger \mathbf{c}_{\mathbf{k}'} \rangle$ is diagonal. From the Hartree-Fock approximation, the Hamiltonian for this system reads

(133)
$$\mathbf{H}^{\text{HF}} = \sum_{\mathbf{k}\sigma} \varepsilon_\mathbf{k}^{\text{HF}} \mathbf{c}_{\mathbf{k}\sigma}^\dagger \mathbf{c}_{\mathbf{k}\sigma}, \text{ where}$$
$$\varepsilon_\mathbf{k}^{\text{HF}} = \varepsilon_\mathbf{k} + \sum_{\mathbf{k}'\sigma'} [V(0) - \delta_{\sigma\sigma'} V(\mathbf{k} - \mathbf{k}')] \mathbf{n}_{\mathbf{k}'\sigma'}$$
$$= \varepsilon_\mathbf{k} + V(0)N - \sum_{\mathbf{k}'\sigma'} V(\mathbf{k} - \mathbf{k}') \mathbf{n}_{\mathbf{k}'\sigma'}.$$

The second term is the interaction with the average electron charge. In condensed matter systems, it is normally cancelled out by an equally large term due to the positively charged ionic background. The third term is precisely the exchange correlation term.

The Hartree-Fock approach depends crucially on what averages are assumed to be finite, and these assumptions must be based on physical knowledge or clever guess-work. For example, in the previous example, spin symmetry is assumed to be maintained, which then implies that $\langle \mathbf{c}_{\mathbf{k}\downarrow} \mathbf{c}_{\mathbf{k}\downarrow} \rangle = \langle \mathbf{c}_{\mathbf{k}\uparrow} \mathbf{c}_{\mathbf{k}\uparrow} \rangle$. Should this not be the case, a ferromagnetic solution may be obtained.

Broken Symmetry. Mean field theory is often used to study phase transitions and thus changes of symmetry. For a given Hamiltonian with some symmetry (eg. translational symmetry, rotational symmetry in real space or in spin space), there exists an operator which reflects this symmetry and therefore commutes with the Hamiltonian (eg. translation operator, rotation operator in either real or spin space). Since the operator and the Hamiltonian commute, a common set of eigenstates exists. Consider for example, the case of a liquid of particles, wherein the Hamiltonian has translational symmetry. Thus, if $\mathcal{T}(\mathbf{R})$ is the translation operator which displaces all particles

coordinates by an $\mathbf{R}$-vector, then $[\mathbf{H}, \mathcal{T}(\mathbf{R})] = 0$. Then, given that this translation operator is in reality a semi-group and given Stone's theorem, then $\mathcal{T} = e^{i\mathbf{R}\cdot\mathbf{P}}$, where P is the total momentum operator. The total momentum operator is then a conserved quantity, given by

$$\mathbf{P} = \hbar \sum_{\mathbf{k}\sigma} \mathbf{k} c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma}.$$

An orthonormal basis of states $|\mathbf{P}\rangle$ with definite total momentum may be chosen. This fact can be used to "prove" the unphysical nature of a density wave. In effect, a $\mathbf{Q}$-density wave can be written in terms of its Fourier components as

$$\rho(\mathbf{Q}) = \sum_{\mathbf{k}\sigma} c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}+\mathbf{Q}\sigma}.$$

Said Forier transform has a finite expectation value. However, it holds as well that

$$\langle c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}+\mathbf{Q}\sigma}\rangle = \frac{1}{\mathcal{Z}} \sum_{\mathbf{P}} e^{-\beta E_{\mathbf{P}}} \langle \mathbf{P}| c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}+\mathbf{Q}\sigma} |\mathbf{P}\rangle = \mathbf{0},$$

since $c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}+\mathbf{Q}\sigma} |\mathbf{P}\rangle$ has momentum $\mathbf{P} - \mathbf{Q}$ and is thus orthogonal to $\mathbf{P}$. Then, according to this result, crystals do not exist. In the same way, magnetism, superconductivity and other physical phenomenon do not exist as well. What is wrong? The previous result breaks down if the sum of states in the thermodynamical average is restricted. Even though crystals with different spatial reference points (or ferromagnets with magnetizations in different directions) have formally the same energy, they are effectively decoupled due to the large energy barrier it takes to melt and then recrystallize into a new state with a shifted distance (or direction of magnetization). In those cases, where many states of the system are degenerate but separated by large energy barriers, it does not make sense to include them on equal footing in the statistical average, since they correspond to macroscopically totally different configurations. Therefore, the ergodicity postulate does not hold, giving rise to a phase space of physically separated sections, which is often illustrated by the double barrier model of phase transitions.

When at some critical temperature, the thermodynamical state of the system develops a non-zero expectation value of some macroscopic quantity, which has less symmetries than the original Hamiltonian, it is called spontaneous symmetry breaking. The quantity which signals that a phase transition has occurred is the order parameter. Some examples of order parameters are presented in table 1.

Examples of Spontaneous Symmetry Breaking		
Phenomena	**Order parameter** **physical**	**Order parameter** **mathematical**
crystal	density wave	$\sum_{\mathbf{k}} \langle c^\dagger_{\mathbf{k}} c_{\mathbf{k}+\mathbf{Q}}\rangle$
ferromagnet	magnetization	$\sum_{\mathbf{k}} \langle c^\dagger_{\mathbf{k}\uparrow} c_{\mathbf{k}\uparrow} - c^\dagger_{\mathbf{k}\downarrow} c_{\mathbf{k}\downarrow}\rangle$
Bose-Einstein condensate	population of $\mathbf{k} = 0$-state	$\langle a^\dagger_{\mathbf{k}=0}\rangle$
superconductor	pair condensate	$\langle c^\dagger_{\mathbf{k}\uparrow} c_{-\mathbf{k}\downarrow}\rangle$

TABLE 1. Typical examples of spontaneous symmetry breakings and their correspondent order parameters.

In order to avoid the paradox obtained before with the wave density, the possibility of phase transitions must be explicitly included into the theory. In the mean field approach, the trick is to include the order parameter in the choice of finite mean fields, and, of course, show that the resulting mean field Hamiltonian leads to a self-consistent finite result.

Ferromagnetism: The Heisenberg model. In ionic magnetic crystals, the interaction between the magnetic ions is due to the exchange interactions originating from the Coulomb interactions. The theory's effective Hamiltonian is given by

$$(134) \qquad \mathbf{H} = -2 \sum_{i\sim j} J_{ij} \mathbf{S}_i \cdot \mathbf{S}, \text{ where } i \sim j \text{ if } ||i - j||_1 = 1.$$

The interaction, which depends only on the inter-ion distance, is generally short ranged and assumed to be simply $J_{ij} = J_0 \delta_{||i-j||_1, 1}$, ie. is only non-zero for nearest-neighbours.

If $J < 0$, the spins tend to become anti-parallel whereas for $J > 0$, it is energetically favourable for the spins to be parallel, the first case corresponds to the antiferromagnetic case while the latter to the ferromagnetic case. For now, consider only the ferromagnetic case, $J > 0$.

As the model stands, it is inherently complicated and cannot be solved in general, the spins of the individual ions being strongly correlated. However, a mean field solution can be readily be found to give sensical physical results. The mean field decomposition reads

(135)
$$\mathbf{H} \approx \mathbf{H}_{\mathrm{MF}} = -2 \sum_{i \sim j} J_{ij} \Big(\langle \mathbf{S}_i \rangle \mathbf{S}_j + \mathbf{S}_i \langle \mathbf{S}_j \rangle - \langle \mathbf{S}_i \rangle \langle \mathbf{S}_j \rangle \Big).$$

From symmetry arguments it is expected that the expectation value of $\langle \mathbf{S}_i \rangle$ is zero, since all directions are equivalent. But this is not the right answer, since a symmetry has been broken, and has to be presumed to be non-zero. Furthermore, since the system has translational symmetry, this average is expected to be position coordinate independent. This would not be the case for an antiferromagnetic solution, wherein position is important since the spins point on opposite directions on even and odd sites. In the ferromagnetic case, a finite but spatially independent average spin polarization is studied. In particular, if the z-axis is chosen to lie along the direction of the magnetization, then the mean field assumption is

$$\langle \mathbf{S}_i \rangle = \langle \mathbf{S}_z \rangle \mathbf{e}_z,$$

and each spin's magnetic moment $\mathbf{m}$, which by assumption is equal for all sites, reads

$$\mathbf{m} = 2 \sum_j J_{ij} \langle \mathbf{S}_z \rangle \mathbf{e}_z = 2 n J_0 \langle \mathbf{S}_z \rangle \mathbf{e}_z, \text{ where } n \text{ is the number of neighbours.}$$

For a square lattice, $\Lambda \subseteq \mathbb{Z}^d$ and $n = 2d$. Then men field Hamiltonian is then

$$\mathbf{H}_{\mathrm{MFT}} = -2 \sum_i \mathbf{m} \cdot \mathbf{S}_i + mN \langle \mathbf{S}_z \rangle,$$

which is now diagonal in the site index and hence easily solved. If the ions are spin-$\frac{1}{2}$ particles, with this simplification, the mean field partition function is

$$\mathcal{Z}_{\mathrm{MFT}} = \Big(e^{\beta m} + e^{-\beta m} \Big)^N e^{\beta N m \langle \mathbf{S}_z \rangle} = \left[\Big(e^{\beta m} + e^{-\beta m} \Big) e^{\frac{\beta m^2}{2n J_0}} \right]^N,$$

with one term for each possible spin projection.

The self-consistency equation is found by minimizing the free energy

$$\frac{\partial F_{\mathrm{MFT}}}{\partial m} = -\frac{1}{\beta} \frac{\partial \log \mathcal{Z}_{\mathrm{MFT}}}{\partial m} = -N \left(\frac{e^{\beta m} - e^{-\beta m}}{e^{\beta m} + e^{-\beta m}} \right) + N \frac{m}{n J_0} = 0 \Rightarrow \alpha = \tanh(b\alpha), \text{ where} \quad \begin{aligned} \alpha &= \frac{m}{n J_0} \\ b &= n J_0 \beta. \end{aligned}$$

For small α, the previous equation reads

$$\alpha \approx b\alpha - \frac{(b\alpha)^3}{3},$$

from which it is evident that there is no solution for $b < 1$. Thus, the critical temperature T_c can be found since at this temperature magnetism disappears. From the $(b_c = 1)$-condition, it is found that $k_B T_c = n J_0$. Furthermore, for small α, the solution for the magnetization can be found

$$(136) \qquad \alpha \approx \frac{1}{b}\sqrt{\frac{3(b-1)}{b}} \Rightarrow m \approx nJ_0\sqrt{3\frac{T_c-T}{T_c}}, \text{ valid for } T \to T_c.$$

At $T=0$, where $t \to \infty$, the solution is $\alpha = 1$ and hence $m = nJ_0$. For the functional form of the magnetization in the entire range of temperature, the transcendental equation must be solved numerically.

The Stoner-Hubbard of metallic ferromagnets. In magnets where the electrons both generate magnetic moments and also form conduction bands, the Heisenberg model proves to be inadequate given that the spins are not localized. Metallic magnetism occurs eg. in transition metals, where the conduction bands are formed by the narrower d or f orbitals. The interaction between two particles in those orbitals is stronger than between electrons occupying the more spread out s or p orbitals and hence give a larger correlation between electrons. Typical metals where correlations between conduction band electrons are important for Fe and Ni.

Since the short-range of the interaction is important, the physical model -the so-called Hubbard model- must incorporate a simple physical fact. Namely, that the Coulomb interaction between electrons is taken to be point-like in real space, and hence constant in momentum space

$$(137) \qquad \mathbf{H}_{\text{Hub}} = \sum_{\mathbf{k}\sigma} \varepsilon_{\mathbf{k}} c_{\mathbf{k}\sigma}^{\dagger} c_{\mathbf{k}\sigma} + \frac{U}{2\mathcal{V}} \sum_{\substack{\mathbf{k},\mathbf{k}',\mathbf{q} \\ \sigma\sigma'}} c_{\mathbf{k}+\mathbf{q}\sigma}^{\dagger} c_{\mathbf{k}'-\mathbf{q}\sigma'}^{\dagger} c_{\mathbf{k}'\sigma'} c_{\mathbf{k}\sigma}.$$

Then, the Hartree-Fock approximation scheme may be employed on this model, by searching for a ferromagnetic solution which allows expectation values to depend on the direction of the spin. The mean field parameters are

$$(138) \qquad \langle c_{\mathbf{k}\uparrow}^{\dagger} c_{\mathbf{k}'\uparrow} \rangle = \delta_{\mathbf{k}\mathbf{k}'}\bar{n}_{\mathbf{k}\uparrow} \qquad\qquad \langle c_{\mathbf{k}\downarrow}^{\dagger} c_{\mathbf{k}'\downarrow} \rangle = \delta_{\mathbf{k}\mathbf{k}'}\bar{n}_{\mathbf{k}\downarrow},$$

and the mean-field interaction Hamiltonian becomes

$$
\begin{aligned}
\mathbf{V}_{\text{int}}^{\text{MFT}} &= \frac{U}{\mathcal{V}} \sum_{\substack{\mathbf{k},\mathbf{k}',\mathbf{q} \\ \sigma\sigma'}} \Bigg[c_{\mathbf{k}+\mathbf{q}\sigma}^{\dagger} \langle c_{\mathbf{k}'-\mathbf{q}\sigma'}^{\dagger} c_{\mathbf{k}'\sigma'} \rangle c_{\mathbf{k}\sigma} - \langle c_{\mathbf{k}+\mathbf{q}\sigma}^{\dagger} c_{\mathbf{k}'\sigma'} \rangle c_{\mathbf{k}'-\mathbf{q}\sigma'}^{\dagger} c_{\mathbf{k}\sigma} \\
&\qquad - \frac{1}{2}\Big(\langle c_{\mathbf{k}+\mathbf{q}\sigma}^{\dagger} c_{\mathbf{k}\sigma} \rangle \langle c_{\mathbf{k}'-\mathbf{q}\sigma'}^{\dagger} c_{\mathbf{k}'\sigma'} \rangle - \langle c_{\mathbf{k}+\mathbf{q}\sigma}^{\dagger} c_{\mathbf{k}'\sigma'} \rangle \langle c_{\mathbf{k}'-\mathbf{q}\sigma'}^{\dagger} c_{\mathbf{k}\sigma} \rangle \Big) \Bigg] \\
&= \frac{U}{\mathcal{V}} \sum_{\substack{\mathbf{k},\mathbf{k}',\mathbf{q} \\ \sigma\sigma'}} \Bigg[c_{\mathbf{k}+\mathbf{q}\sigma}^{\dagger} c_{\mathbf{k}\sigma} \delta_{\mathbf{k}'-\mathbf{q}\,\mathbf{k}'} \bar{n}_{\mathbf{k}'\sigma'} - c_{\mathbf{k}'-\mathbf{q}\sigma'}^{\dagger} c_{\mathbf{k}\sigma} \delta_{\mathbf{k}+\mathbf{q}\,\mathbf{k}'} \delta_{\sigma\sigma'} \bar{n}_{\mathbf{k}'\sigma'} \\
&\qquad - \frac{1}{2}\Big(\delta_{\mathbf{k}+\mathbf{q}\,\mathbf{k}} \bar{n}_{\mathbf{k}\sigma} \delta_{\mathbf{k}'-\mathbf{q}\,\mathbf{k}'} \bar{n}_{\mathbf{k}'\sigma'} - \delta_{\mathbf{k}+\mathbf{q}\,\mathbf{k}'} \delta_{\sigma\sigma'} \bar{n}_{\mathbf{k}'\sigma'} \delta_{\mathbf{k}'-\mathbf{q}\,\mathbf{k}} \delta_{\sigma\sigma'} \bar{n}_{\mathbf{k}\sigma} \Big) \Bigg] \\
&= U \sum_{\substack{\mathbf{k},\mathbf{k}' \\ \sigma\sigma'}} \Bigg[c_{\mathbf{k}\sigma}^{\dagger} c_{\mathbf{k}\sigma} \frac{\bar{n}_{\mathbf{k}'\sigma'}}{\mathcal{V}} - c_{\mathbf{k}\sigma}^{\dagger} c_{\mathbf{k}\sigma} \delta_{\sigma\sigma'} \frac{\bar{n}_{\mathbf{k}'\sigma'}}{\mathcal{V}} - \frac{1}{2}\Big(\bar{n}_{\mathbf{k}\sigma} \frac{\bar{n}_{\mathbf{k}'\sigma'}}{\mathcal{V}} - \delta_{\sigma\sigma'} \bar{n}_{\mathbf{k}'\sigma'} \frac{\bar{n}_{\mathbf{k}\sigma}}{\mathcal{V}} \Big) \Bigg] \\
&= U \sum_{\mathbf{k}\sigma\sigma'} \Bigg[c_{\mathbf{k}\sigma}^{\dagger} c_{\mathbf{k}\sigma} (\bar{n}_{\sigma'} - \delta_{\sigma\sigma'} \bar{n}_{\sigma}) - \mathcal{V}\bar{n}_{\sigma}\bar{n}_{\sigma'} + \mathcal{V}\bar{n}_{\sigma}^2 \Bigg], \text{ where } \bar{n}_{\sigma} = \frac{1}{\mathcal{V}} \sum_{\mathbf{k}'} \langle c_{\mathbf{k}'\sigma}^{\dagger} c_{\mathbf{k}'\sigma} \rangle \text{ is the spin density.}
\end{aligned}
$$

The full Hamiltonian reads

$$(139) \quad \mathbf{H}_{\text{MFT}} = \sum_{\mathbf{k}\sigma} \varepsilon_{\mathbf{k}\sigma}^{\text{MFT}} c_{\mathbf{k}\sigma}^{\dagger} c_{\mathbf{k}} - \frac{U\mathcal{V}}{2} \sum_{\sigma\sigma'} \Big(\bar{n}_{\sigma}\bar{n}_{\sigma'} - \bar{n}_{\sigma}^2 \delta_{\sigma\sigma'} \Big), \text{ with } \varepsilon_{\mathbf{k}\sigma}^{\text{MFT}} = \varepsilon_{\mathbf{k}} + U(n_{\uparrow} + \bar{n}_{\downarrow} - \bar{n}_{\sigma}) = \varepsilon_{\mathbf{k}} + U\bar{n}_{\bar{\sigma}}.$$

The mean field solution is found via minimization,

$$0 = \frac{d}{d\bar{n}_\sigma} F_{\mathrm{MFT}} = \frac{d}{d\bar{n}_\sigma}\Big(-\frac{\log \mathcal{Z}_{\mathrm{MFT}}}{\beta}\Big) = \frac{1}{\mathcal{Z}_{\mathrm{MFT}}}\,\mathrm{Tr}\Big(e^{-\beta \mathbf{H}_{\mathrm{MFT}}}\frac{d}{d\bar{n}_\sigma}\mathbf{H}_{\mathrm{MFT}}\Big)$$

$$= -\frac{1}{\mathcal{Z}_{\mathrm{MFT}}}\,\mathrm{Tr}\Big[e^{-\beta \mathbf{H}_{\mathrm{MFT}}}\Big(\sum_{\mathbf{k}\sigma}\frac{d\varepsilon^{\mathrm{MFT}}_{\mathbf{k}\sigma}}{d\bar{n}_\sigma}\mathbf{c}^\dagger_{\mathbf{k}\sigma}\mathbf{c}_{\mathbf{k}\sigma} - \frac{U\mathcal{V}}{2}\sum_{\sigma\sigma'}(\bar{n}_{\sigma'}+\bar{n}_\sigma\delta_{\sigma\sigma'}-2\bar{n}_{\sigma'}\delta_{\sigma\sigma'})\Big)\Big]$$

$$= -\frac{1}{\mathcal{Z}_{\mathrm{MFT}}}\,\mathrm{Tr}\Big[e^{\beta \mathbf{H}_{\mathrm{MFT}}}\Big(\sum_{\mathbf{k}\sigma}-U\mathbf{c}^\dagger_{\mathbf{k}\sigma}\mathbf{c}_{\mathbf{k}\sigma} - \frac{U\mathcal{V}}{2}\sum_{\sigma\sigma'}(\bar{n}_{\sigma'}-\bar{n}_{\sigma'}\delta_{\sigma\sigma'})\Big)\Big]$$

$$= U\sum_{\mathbf{k}\sigma}\langle \mathbf{c}^\dagger_{\mathbf{k}\sigma}\mathbf{c}_{\mathbf{k}\sigma}\rangle_{\mathrm{MFT}} - \frac{U\mathcal{V}}{2}\sum_{\sigma\sigma'}(\bar{n}_{\sigma'}-\bar{n}_{\sigma'}\delta_{\sigma\sigma'}) = U\langle \mathbf{c}^\dagger_{\mathbf{k}\sigma}\mathbf{c}_{\mathbf{k}}\rangle_{\mathrm{MFT}} + U\mathcal{V}\bar{n}_\sigma$$

$$= \frac{U}{\mathcal{Z}_{\mathrm{MFT}}}\sum_{\mathbf{k}\sigma}\mathrm{Tr}\Big[e^{-\beta \mathbf{H}_{\mathrm{MFT}}}\mathbf{c}^\dagger_{\mathbf{k}\sigma}\mathbf{c}_{\mathbf{k}\sigma}\Big] + \frac{U\mathcal{V}}{2}\sum_{\sigma\sigma'}(\bar{n}_{\sigma'}-\bar{n}_{\sigma'}\delta_{\sigma\sigma'})\frac{1}{\mathcal{Z}_{\mathrm{MFT}}}\mathrm{Tr}\Big[\cancel{e^{-\beta \mathbf{H}_{\mathrm{MFT}}}\mathbb{1}}\Big]$$

$$= \frac{U}{\mathcal{V}}\sum_{\mathbf{k}}\langle \mathbf{c}^\dagger_{\mathbf{k}\sigma}\mathbf{c}_{\mathbf{k}\sigma}\rangle_{\mathrm{MFT}} - \frac{U\mathcal{V}}{2}\Big(2\sum_{\sigma'}\bar{n}_{\sigma'} - \sum_\sigma \bar{n}_\sigma\Big)\Big] = \sum_\sigma\Big[\sum_{\mathbf{k}}\frac{U}{\mathcal{V}}\langle \mathbf{c}^\dagger_{\mathbf{k}\sigma}\mathbf{c}_{\mathbf{k}\sigma}\rangle_{\mathrm{MFT}} + \frac{U\mathcal{V}}{2}\bar{n}_\sigma\Big]$$

$$\Rightarrow \bar{n}_\sigma = \frac{1}{\mathcal{V}}\sum_{\mathbf{k}}\langle \mathbf{c}^\dagger_{\mathbf{k}\sigma}\mathbf{c}_{\mathbf{k}\sigma}\rangle_{\mathrm{MFT}} = \frac{1}{\mathcal{V}}\sum_{\mathbf{k}}n_F(\varepsilon^{\mathrm{MFT}}_{\mathbf{k}\sigma}).$$

At zero temperature,

(140)
$$\bar{n}_\uparrow = \int_{\mathbb{R}^3}\frac{d\mathbf{k}}{(2\pi)^3}\,\Theta\Big(\mu - \frac{\hbar^2 k^2}{2m} - U\bar{n}_\downarrow\Big) = \frac{k^3_{F\uparrow}}{6\pi^2}$$
$$\bar{n}_\downarrow = \int_{\mathbb{R}^3}\frac{d\mathbf{k}}{(2\pi)^3}\,\Theta\Big(\mu - \frac{\hbar^2 k^2}{2m} - U\bar{n}_\uparrow\Big) = \frac{k^3_{F\downarrow}}{6\pi^2},$$
where $\frac{\hbar^2 k_{F\sigma}}{2m} + U\bar{n}_{(\sigma)^*} = \mu.$

Thus, yielding two equations

(141)
$$\frac{\hbar^2}{2m}(6\pi)^{\frac{2}{3}}\bar{n}_\uparrow^{\frac{2}{3}} + U\bar{n}_\downarrow = \mu, s \qquad\qquad \frac{\hbar^2}{2m}(6\pi)^{\frac{2}{3}}\bar{n}_\downarrow^{\frac{2}{3}} + U\bar{n}_\uparrow = \mu.$$

Let,

$$\zeta = \frac{\bar{n}_\uparrow - \bar{n}_\downarrow}{\bar{n}}, \qquad\qquad \gamma = \frac{2mU n^{\frac{1}{3}}}{(3\pi^2)^{\frac{3}{2}}\hbar^2}, \bar{n} \qquad\qquad = \bar{n}_\uparrow + \bar{n}_\downarrow,$$

then, by subtracting the two equations,

(142)
$$\bar{n}_\uparrow^{\frac{2}{3}} - \bar{n}_\downarrow^{\frac{2}{3}} = \frac{2mU}{\hbar^2}(6\pi)^{-\frac{2}{3}}(\bar{n}_\uparrow - \bar{n}_\downarrow) \Rightarrow \gamma\zeta = (1+\zeta)^{\frac{2}{3}} - (1-\zeta)^{\frac{2}{3}},$$

which has three types of solutions:

- $\gamma < \frac{4}{3}$: isotropic solution: $\zeta = 0$, which corresponds to a normal state.
- $\frac{4}{3} < \gamma < 2^{\frac{2}{3}}$: Partial polarization $0 < \zeta < 1$, which corresponds to a weak ferromagnet.
- $\gamma > 2^{\frac{2}{3}}$: Full polarization $\zeta = 1$, which corresponds to a strong ferromagnet.

The possibility for a magnetic solution can be traced back to the spin dependent eigenenergies, where it is clear that the mean field energy of a given spin direction depends on the occupation of the opposite spin direction, whereas the energy does not depend on the occupation of the same spin direction. This is a consequence of two things: the short range interaction and the exchange term. Both can be understood from the Pauli principle, which ensures that electrons with the same spin never occupy the same spatial orbitals and therefore, if the interaction is short-range, they cannot interact. This leaves interactions between opposite spins as the only possibility. Thus, the interaction energy is lowered by having a polarized ground state, which on the other hand for a fixed density costs kinetic energy. The competition between the kinetic and potential energy gives rise to the phase transition.

The Stoner-Hubbard model gives a reasonable account for metallic magnets and is capable of qualitatively explaining the properties of excitations in the spin polarized states.

Superconductivity: BCS theory. One of the most striking examples of symmetry breaking is the superconducting phase transition. Below the critical temperature the metal which turns superconducting, has no resistance whatsoever, and it exhibits perfect diamagnetism, called Meisner effect, which means that magnetic fields are totally expelled from the interior of the material. These astonishing phenomena result from the superconducting state having a new form of symmetry breaking, namely loss of global gauge invariance. Both the new type of phase and the appearance of an excitation gap are explained by the Bardeen-Cooper-Schrieffer theory.

<u>Breaking of global gauge symmetry and its consequences.</u> What kind of broken symmetry may give rise to a superconducting state? the relevant symmetry is the global gauge symmetry, ie. all electrons may be added the same constant extra phase without altering the Hamiltonian. The ferromagnetic analog is that all spins can be rotated by some angle without changing the Hamiltonian, due to a global $SO(3)$-symmetry. In that case, a broken $SO(3)$-symmetry, entails that the expectation value $\langle \mathbf{S} \rangle$ is not invariant under rotation, since it will change the magnetization's direction. In the same venue, a phase rotation changes the superconducting order parameter, which is of the form $\langle \mathbf{c}_\mu \mathbf{c}_{\mu'} \rangle$ ie. the expectation value of two annihilation operators. Of course, the number of particles is conserved, but this is not a problem in our setting for we are only interested in superconductors connected to electron reservoirs.

While it is clear why the finite expectation value $\langle \mathbf{S} \rangle$ gives a magnetization in the case of a ferromagnet, it is not so clear why broken symmetry in the superconducting case leads to a system without resistance. Moreover, even though the superconducting phase is sensitive to a change in global phase, it is clear that a constant phase cannot have any measurable effect, since all expectation values are given in terms of the absolute square of the wave function. However, phase gradients can have an effect.

Consider a superconducting state, which is ascribed by a phase which depends on the position, $\phi(\mathbf{x})$, and varies extremely slowly, such that it takes a macroscopic distance to see any significant changes in $\phi(\mathbf{x})$. For any other non-superconducting system, it would not make sense to talk about a quantum mechanical phase difference over macroscopic distances, simply because quantum coherence is destroyed by all sorts of scattering events on rather short lengths ($\sim 10nm$ for metals). To argue that the superconducting state depends on phase differences over macroscopic distances is unusual and implies that superconductivity is a macroscopic quantum phenomenon.

Consider the following unitary transformation, which changes the phase

$$U = \exp\left(i \int_{\mathbb{R}^3} d\mathbf{x}\, \rho(\mathbf{x})\phi(\mathbf{x})\right), \text{ where } \rho(\mathbf{x}) = \Psi^\dagger(\mathbf{x})\Psi(\mathbf{x}).$$

When applied to quantum field operators it has the following properties

(143) $$\tilde{\Psi}(\mathbf{x}) = U\Psi(\mathbf{x})U^{-1} = \exp\left(i \int_{\mathbb{R}^3} d\mathbf{x}\, \rho(\mathbf{x})\phi(\mathbf{x})\right)\Psi(\mathbf{x})\exp\left(-i \int_{\mathbb{R}^3} d\mathbf{x}\, \rho(\mathbf{x})\phi(\mathbf{x})\right)$$

(144) $$= \Psi(\mathbf{x})e^{-i\phi(\mathbf{x})}$$

(145) $$\tilde{\Psi}^\dagger(\mathbf{x}) = U\Psi^\dagger(\mathbf{x})U^{-1} = \Psi^\dagger(\mathbf{x})e^{i\phi(\mathbf{x})}.$$

These equations follow from the differential equation problem

$$\begin{cases} \frac{\delta\tilde{\Psi}(\mathbf{x})}{\delta\phi(\mathbf{x}')} = iU[\rho(\mathbf{x}'), \Psi(\mathbf{x})]U^{-1} = i\tilde{\Psi}(\mathbf{x})\delta(\mathbf{x} - \mathbf{x}') \\ \left.\tilde{\Psi}(\mathbf{x})\right|_{\phi=0} = \Psi(\mathbf{x}) \end{cases}$$

The transformed partition function is given by

$$\tilde{\mathcal{Z}} = \text{Tr}'[Ue^{-\beta\mathbf{H}}U^{-1}] = \text{Tr}'[e^{-\beta\tilde{\mathbf{H}}}],$$

note that, if the cyclic property of the trace had been used, U would have disappeared all-together. However, when dealing with systems with broken symmetry, the sum-over-states has to be restricted so that the the cyclic property does not hold. In other words, the trace is restricted due to the spontaneous symmetry breaking. The transformation only changes the kinetic term of the Hamiltonian, since the Coulomb interaction term, the impurity scattering term and the phonon coupling term only involves the, phase-invariant, electron density operator $\rho(\mathbf{x})$. The kinetic term reads

(146)
$$\begin{aligned}
\tilde{\mathbf{H}} &= \frac{1}{2m_e} \int_{\mathbb{R}^3} d\mathbf{x}\ \boldsymbol{\Psi}^\dagger(\mathbf{x}) e^{i\phi(\mathbf{x})} \left(\frac{\hbar}{i}\nabla + e\mathbf{A}\right)^2 e^{-i\phi(\mathbf{x})} \boldsymbol{\Psi}(\mathbf{x}) \\
&= \frac{1}{2m_e} \int_{\mathbb{R}^3} d\mathbf{x}\ \boldsymbol{\Psi}^\dagger(\mathbf{x}) \left(\frac{\hbar}{i}\nabla + e\mathbf{A} - \hbar\nabla\phi(\mathbf{x})\right)^2 \boldsymbol{\Psi}(\mathbf{x}) \\
&= \mathbf{H} - \hbar \int_{\mathbb{R}^3} d\mathbf{x}\ \nabla\phi(\mathbf{x})\cdot\mathbf{J}(\mathbf{x}) - \frac{\hbar}{2m_e}\rho(\mathbf{x})(\nabla\phi(\mathbf{x}))^2.
\end{aligned}$$

Therefore, the phase is a macroscopic quantity. The lowest free energy condition can be found by minimizing the free energy with respect to the phase. It is clear that the energy does not depend with ϕ itself, but on its gradient, hence

(147)
$$\begin{aligned}
\frac{\delta F}{\delta\phi(\mathbf{x})} &= -\hbar \int_{\mathbb{R}^3} d\mathbf{x}\ \mathbf{J}(\mathbf{x}) - \frac{\hbar}{m_e}\rho(\mathbf{x})\nabla\phi(\mathbf{x}) \\
&= -\langle\mathbf{J}(\mathbf{x})\rangle + \frac{\hbar}{m_e}\langle\rho(\mathbf{x})\rangle\nabla\phi(\mathbf{x}) = 0 \Rightarrow \langle\mathbf{J}\rangle = \frac{\hbar\rho_0}{m_e}\nabla\phi.
\end{aligned}$$

The previous equation's physical meaning is that, by forcing a phase gradient onto the system, it minimizes its energy by carrying a current even in thermodynamical equilibrium, giving rise to a non-equilibrium current. In reality, the electron density should only include those electrons participating in the superconducting condensate, the so-called superfluid. In this derivation, it is assumed all electrons participate.

In the normal state of metals, a current is always associated with a non-equilibrium state, where energy is constantly dissipated from the driving source and absorbed in the conductor. Of course, there is an energy cost for the system to carry the current, but as long as this cost is smaller than the alternative which is to go out of the superconducting state, the current carrying state is chosen. The critical current is reaches when the energies are equal, and then the superconductor goes into the normal state. Thus, when a phase gradient is imposed to a system, wherein the energy is assumed to depend on phase differences on a macroscopic scale, it unavoidably leads to the conclusion that the system will carry a dissipationless current in order to minimize the energy cost of the phase gradient. Finally, it should be noted that the appearance of the excitation gap is not the reason for the superconductivity itself. The superconductivity arises from the lack of gauge invariance, and in fact gapless superconductors do exist.

Microscopic theory:

Historical context: the understanding that superconductivity was closely related to the electron-phonon coupling was clear from the 1950's onwards, where the isotope effect was discovered. Also the idea that superconductivity was somehow similar to the Bose-Einstein condensate, with the bosons being electron pairs, had been tried and in fact was the underlying idea of London's theory in 1935. However, only in 1956 Cooper showed that the Fermi surface of the normal mental state was unstable towards formation of bound pairs of electrons. Subsequently in 1957, when the superconducting state was derived using a variation wavefunction by BCS were the principles fully understood. It turns out that the phonon mediante electron-electron interaction, in fact has a range in frequency and momentum space where it is negative, ie. attracitive. This happens for energy exchanges of order the Debye energy, ω_D, for metals is much smaller than the Fermi energy $\omega_D << E_F$. Furthermore, from the Cooper instability, it follows that the photon-mediated interactions tends to pair electrons with opposite spin and momentum.

Let a Cooper pair operator be

(148)
$$\mathbf{b}_k = \mathbf{c}_{k\downarrow}\mathbf{c}_{-k\uparrow}.$$

Then, the previous two physical inputs, led BCS to suggest the following remarkably successful model Hamiltonian to explain the superconducting state. The BCS effective Hamiltonian model is

(149)
$$\mathbf{H}_{\text{BCS}} = \sum_{k\sigma} \varepsilon_\mathbf{k} \mathbf{c}_{\mathbf{k}\sigma}^\dagger \mathbf{c}_{\mathbf{k}\sigma} + \sum_{\mathbf{k}\mathbf{k}'} \mathfrak{V}_{\mathbf{k}\mathbf{k}'} \mathbf{c}_{\mathbf{k}\uparrow}^\dagger \mathbf{c}_{-\mathbf{k}\downarrow}^\dagger \mathbf{c}_{-\mathbf{k}'\downarrow} \mathbf{c}_{\mathbf{k}',\uparrow},$$

where $\mathfrak{V}_{\mathbf{k}\mathbf{k}'} = -V_0 \Big(\Theta(\varepsilon_\mathbf{k} + \epsilon_F + \omega_D) - \Theta(\varepsilon_\mathbf{k} - \epsilon_F - \omega_D) \Big) \delta_{\mathbf{k}\mathbf{k}'} \neq 0$ if $\varepsilon_\mathbf{k} \in (\epsilon_F - \omega_D, \epsilon_F + \omega_D)$. The interaction includes only pair interactions and the remaining interaction is supposed to be included in $\varepsilon_\mathbf{k}$ via a Hartree-Fock term. The attractive interaction is due to the phenomenon of an electron propagating through a crystal, thereby attracting the positive ions and thus effectively creating a positive trace behind it, which is felt by the other electrons as an attractive interaction. It turns out that this effective interaction is most important for electrons occupying time reversed states, and in fact they can form a bounded state which is a Cooper pair. The Cooper pairs is thus a bounded state of an electron in state $\psi_\mu(\mathbf{x})$ and another electron in state $\psi_\mu^*(\mathbf{x})$, or in the homogeneous case, electrons in states $\mathbf{k}$ and $-\mathbf{k}$. [15]

The mean field assumption made by BCS is that the pair operator has a finite expectation value, which deviates only a little from its average values. Therefore, the BCS mean field Hamiltonian is derived in full analogy with Hartree-Fock's mean field theory, yielding

(150)
$$\begin{aligned} \mathbf{H}_{\text{BCS}}^{\text{MFT}} = {} & \sum_{k\sigma} \varepsilon_\mathbf{k} \mathbf{c}_{\mathbf{k}\sigma}^\dagger \mathbf{c}_{\mathbf{k}\sigma} - \sum_\mathbf{k} \Delta_\mathbf{k} \mathbf{c}_{\mathbf{k}\uparrow}^\dagger \mathbf{c}_{-\mathbf{k}\downarrow}^\dagger - \sum_\mathbf{k} \Delta_\mathbf{k}^* \mathbf{c}_{\mathbf{k}\downarrow} \mathbf{c}_{-\mathbf{k}\uparrow} \\ & + \sum_{\mathbf{k}\mathbf{k}'} \mathfrak{V}_{\mathbf{k}\mathbf{k}'} \langle \mathbf{c}_{\mathbf{k}\uparrow}^\dagger \mathbf{c}_{-\mathbf{k}\downarrow}^\dagger \rangle \langle \mathbf{c}_{\mathbf{k}\downarrow} \mathbf{c}_{-\mathbf{k}\uparrow} \rangle, \text{ where } \Delta_\mathbf{k} = -\sum_{\mathbf{k}'} \mathfrak{V}_{\mathbf{k}\mathbf{k}'} \langle \mathbf{c}_{-\mathbf{k}'\downarrow} \mathbf{c}_{\mathbf{k}'\uparrow} \rangle, \end{aligned}$$

which is quadratic in electron operators and should be readily solvable. It is however unusual in that terms $\mathbf{c}^\dagger \mathbf{c}^\dagger$ and $\mathbf{c}\mathbf{c}$ both appear. Said Hamiltonian can then be solved via a Bogoliubov transformation. Note that $\mathbf{H}_{\text{BCS}}^{\text{MFT}}$ may be rewritten in $\mathbb{C}^2$-matrix form

(151)
$$\mathbf{H}_{\text{BCS}}^{\text{MFT}} = \sum_\mathbf{k} \begin{pmatrix} \mathbf{c}_{\mathbf{k}\uparrow}^\dagger & \mathbf{c}_{-\mathbf{k}\downarrow} \end{pmatrix} \begin{pmatrix} \varepsilon_\mathbf{k} & \Delta_\mathbf{k} \\ \Delta_\mathbf{k}^* & -\varepsilon_\mathbf{k} \end{pmatrix} \begin{pmatrix} \mathbf{c}_{\mathbf{k}\downarrow}^\dagger \\ \mathbf{c}_{-\mathbf{k}\uparrow} \end{pmatrix} + \sum_\mathbf{k} \varepsilon_\mathbf{k} + \sum_{\mathbf{k}\mathbf{k}'} \mathfrak{V}_{\mathbf{k}\mathbf{k}'} \langle \mathbf{c}_{\mathbf{k}\uparrow}^\dagger \mathbf{c}_{-\mathbf{k}\downarrow}^\dagger \rangle \langle \mathbf{c}_{\mathbf{k}\downarrow} \mathbf{c}_{-\mathbf{k}\uparrow} \rangle$$

(152)
$$= \sum_\mathbf{k} \mathbf{A}_\mathbf{k}^\dagger \mathbf{H}_k \mathbf{A}_\mathbf{k} + \text{tr}\{\mathbf{H}\},$$

where

$$\mathbf{A}_\mathbf{k} = \begin{pmatrix} \mathbf{c}_{\mathbf{k}\uparrow} \\ \mathbf{c}_{-\mathbf{k}\downarrow}^\dagger \end{pmatrix}.$$

Then, the Hamiltonian can be rewritten in diagonal for, via a redefinition of operators, said redefinition being given by a unitary transformation,

(153)
$$\mathbf{B}_\mathbf{k} = \begin{pmatrix} \gamma_{\mathbf{k}\uparrow} \\ \gamma_{-\mathbf{k}\downarrow}^\dagger \end{pmatrix} = \mathbf{U}_\mathbf{k}^{-1} \mathbf{A}_\mathbf{k} \text{ where } \mathbf{U}_\mathbf{k} = \begin{pmatrix} u_\mathbf{k} & -v_\mathbf{k} \\ v_\mathbf{k}^* & u_\mathbf{k}^* \end{pmatrix},$$

---

[15]Note that, since the typical energy exchange due to the attractive interaction is the Debye energy ω_D one could naively expect that the energy scale for the superconducting transition temperature would be of $\mathcal{O}(\frac{\omega_D}{k})$-order. This is however far from the truth, since $\frac{\omega}{k}$ is typically of the order of several hundred kelvin, while the critical temperatures found in conventional low superconductors are, in most cases, less than 10K. Thus, it seems that a new energy scale is needed.

which diagonalizes the Hamiltonian provided

(154)

$$\mathbf{U}_\mathbf{k}^\dagger \mathbf{H}_\mathbf{k} \mathbf{U}_\mathbf{k} = \begin{pmatrix} u_\mathbf{k}^* & v_\mathbf{k} \\ -v_\mathbf{k}^* & u_\mathbf{k} \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} u_\mathbf{k}^* & -v_\mathbf{k}^* \\ v_\mathbf{k} & u_\mathbf{k}^* \end{pmatrix} \begin{pmatrix} \varepsilon_\mathbf{k} & \Delta_\mathbf{k} \\ \Delta_\mathbf{k}^* & -\varepsilon_\mathbf{k} \end{pmatrix} \begin{pmatrix} u_\mathbf{k} & -v_\mathbf{k} \\ v_\mathbf{k}^* & u_\mathbf{k}^* \end{pmatrix} = \begin{pmatrix} u_\mathbf{k}^* & v_\mathbf{k} \\ -v_\mathbf{k}^* & u_\mathbf{k} \end{pmatrix} \begin{pmatrix} \varepsilon_\mathbf{k} u_\mathbf{k} + \Delta_\mathbf{k} v_\mathbf{k}^* & -\varepsilon_\mathbf{k} v_\mathbf{k} + \Delta_\mathbf{k} u_\mathbf{k}^* \\ \Delta_\mathbf{k}^* u_\mathbf{k} - \varepsilon_\mathbf{k} v_\mathbf{k}^* & -\Delta_\mathbf{k}^* v_\mathbf{k} - \varepsilon_\mathbf{k} u_\mathbf{k}^* \end{pmatrix}$$

$$= \begin{pmatrix} \varepsilon_\mathbf{k}|u_\mathbf{k}|^2 + \Delta_\mathbf{k} u_\mathbf{k}^* v_\mathbf{k}^* + \Delta_\mathbf{k}^* u_\mathbf{k} v_\mathbf{k} - \varepsilon_\mathbf{k}|v_\mathbf{k}|^2 & -\varepsilon_\mathbf{k} u_\mathbf{k}^* v_\mathbf{k} + \Delta_\mathbf{k}|u_\mathbf{k}|^2 - \Delta_\mathbf{k}^* v_\mathbf{k}^2 - \varepsilon_\mathbf{k} u_\mathbf{k}^* v_\mathbf{k} \\ -\varepsilon_\mathbf{k} u_\mathbf{k} v_\mathbf{k}^* - \Delta_\mathbf{k}|v_\mathbf{k}|^2 + \Delta_\mathbf{k}^* u_\mathbf{k}^2 - \varepsilon_\mathbf{k} u_\mathbf{k} v_\mathbf{k}^* & \varepsilon_\mathbf{k}|v_\mathbf{k}|^2 - \Delta_\mathbf{k} u_\mathbf{k}^* v_\mathbf{k}^* - \Delta_\mathbf{k}^* v_\mathbf{k} u_\mathbf{k} - \varepsilon_\mathbf{k}|u_\mathbf{k}|^2 \end{pmatrix}$$

$$\Rightarrow \begin{matrix} |u_\mathbf{k}|^2 = \frac{1}{2}\left(1 + \frac{\varepsilon_\mathbf{k}}{E_\mathbf{k}}\right) \\ |v_\mathbf{k}|^2 = \frac{1}{2}\left(1 - \frac{\varepsilon_\mathbf{k}}{E_\mathbf{k}}\right) \end{matrix} \quad E_\mathbf{k} = \sqrt{\varepsilon_\mathbf{k}^2 + |\Delta_\mathbf{k}|^2} = -\tilde{E}_\mathbf{k}.$$

The new fermionic operators, which diagonalize the Hamiltonian, are fermionic quasiparticles, frequently called bogoliubons and are the superposition of electrons and holes. This new particle arises due to a lack of particle conservation in the mean field Hamiltonian. Note that there are two different bogoliubons inherited from the two-fold spin degeneracy. The quasiparticles operators are given by

$$\begin{pmatrix} \gamma_{\mathbf{k}\uparrow} \\ \gamma_{-\mathbf{k}\downarrow}^\dagger \end{pmatrix} = \begin{pmatrix} u_\mathbf{k}^* & v_\mathbf{k} \\ -v_\mathbf{k}^* & u_\mathbf{k} \end{pmatrix} \begin{pmatrix} c_{\mathbf{k}\uparrow} \\ c_{-\mathbf{k}\downarrow}^\dagger \end{pmatrix} \Leftrightarrow \begin{pmatrix} c_{\mathbf{k}\uparrow} \\ c_{-\mathbf{k}\downarrow}^\dagger \end{pmatrix} = \begin{pmatrix} u_\mathbf{k} & -v_\mathbf{k} \\ v_\mathbf{k}^* & u_\mathbf{k}^* \end{pmatrix} \begin{pmatrix} \gamma_{\mathbf{k}\uparrow} \\ \gamma_{-\mathbf{k}\downarrow}^\dagger \end{pmatrix}$$

(155)

$$\begin{pmatrix} \gamma_{\mathbf{k}\uparrow} \\ \gamma_{-\mathbf{k}\downarrow}^\dagger \\ \gamma_{\mathbf{k}\uparrow}^\dagger \\ \gamma_{-\mathbf{k}\downarrow} \end{pmatrix} = \begin{pmatrix} u_\mathbf{k}^* c_{\mathbf{k}\uparrow} + v_\mathbf{k} c_{-\mathbf{k}\downarrow}^\dagger \\ -v_\mathbf{k}^* c_{\mathbf{k}\uparrow} + u_\mathbf{k} c_{-\mathbf{k}\downarrow}^\dagger \\ u_\mathbf{k} c_{\mathbf{k}\uparrow}^\dagger + v_\mathbf{k}^* c_{-\mathbf{k}\downarrow} \\ -v_\mathbf{k} c_{\mathbf{k}\uparrow}^\dagger + u_\mathbf{k} c_{-\mathbf{k}\downarrow} \end{pmatrix} \Leftrightarrow \begin{pmatrix} c_{\mathbf{k}\uparrow} \\ c_{-\mathbf{k}\downarrow}^\dagger \\ c_{\mathbf{k}\uparrow}^\dagger \\ c_{-\mathbf{k}\downarrow} \end{pmatrix} = \begin{pmatrix} u_\mathbf{k} \gamma_{\mathbf{k}\uparrow} - v_\mathbf{k} \gamma_{-\mathbf{k}\downarrow}^\dagger \\ v_\mathbf{k}^* \gamma_{\mathbf{k}\uparrow} + u_\mathbf{k}^* \gamma_{-\mathbf{k}\downarrow}^\dagger \\ u_\mathbf{k}^* \gamma_{\mathbf{k}\uparrow}^\dagger - v_\mathbf{k}^* \gamma_{-\mathbf{k}\downarrow} \\ v_\mathbf{k} \gamma_{\mathbf{k}\uparrow}^\dagger + u_\mathbf{k} \gamma_{-\mathbf{k}\downarrow} \end{pmatrix}.$$

It is necessary to inquire in the bogoliubons' nature. Namely, it they obey fermionic anticommutation rules,

$$\{\gamma_{\mathbf{k}\uparrow}^\dagger, \gamma_{\mathbf{k}'\uparrow}\} = \{u_\mathbf{k} c_{\mathbf{k}\uparrow}^\dagger + v_\mathbf{k}^* c_{-\mathbf{k}\downarrow}, u_{\mathbf{k}'}^* c_{\mathbf{k}'\uparrow} + v_{\mathbf{k}'} c_{-\mathbf{k}'\downarrow}^\dagger\}$$
$$= u_\mathbf{k} u_{\mathbf{k}'}^* \{c_{\mathbf{k}\uparrow}^\dagger, c_{\mathbf{k}'\uparrow}\} + u_\mathbf{k}^* v_{\mathbf{k}'}^* \cancel{\{c_{-\mathbf{k}\downarrow}, c_{\mathbf{k}'\uparrow}\}}$$
$$+ u_{\mathbf{k}'} v_\mathbf{k} \cancel{\{c_{\mathbf{k}\uparrow}^\dagger, c_{-\mathbf{k}'\downarrow}^\dagger\}} + v_\mathbf{k}^* v_{\mathbf{k}'} \{c_{-\mathbf{k}\downarrow}, c_{-\mathbf{k}'\downarrow}^\dagger\}$$
$$= u_\mathbf{k} u_{\mathbf{k}'}^* \delta_{\mathbf{k}\mathbf{k}'} + v_\mathbf{k} v_{\mathbf{k}'}^* \delta_{\mathbf{k}\mathbf{k}'} = \delta_{\mathbf{k}\mathbf{k}'}$$

$$\{\gamma_{-\mathbf{k}\downarrow}^\dagger, \gamma_{-\mathbf{k}'\downarrow}\} = \{-v_\mathbf{k}^* c_{\mathbf{k}\uparrow} + u_\mathbf{k} c_{-\mathbf{k}\downarrow}^\dagger, -v_{\mathbf{k}'} c_{\mathbf{k}'\uparrow}^\dagger + u_{\mathbf{k}'} c_{-\mathbf{k}'\downarrow}\}$$
$$= v_\mathbf{k} v_{\mathbf{k}'}^* \{c_{\mathbf{k}\uparrow}, c_{\mathbf{k}'\uparrow}^\dagger\} - u_\mathbf{k} v_{\mathbf{k}'} \cancel{\{c_{-\mathbf{k}\downarrow}^\dagger, c_{\mathbf{k}'\uparrow}^\dagger\}}$$
$$- u_{\mathbf{k}'} v_\mathbf{k}^* \cancel{\{c_{-\mathbf{k}\downarrow}^\dagger, c_{\mathbf{k}'\uparrow}^\dagger\}} + v_\mathbf{k} v_{\mathbf{k}'}^* \{c_{-\mathbf{k}\downarrow}^\dagger, c_{-\mathbf{k}'\downarrow}\}$$
$$= u_\mathbf{k} u_{\mathbf{k}'}^* \delta_{\mathbf{k}\mathbf{k}'} + v_\mathbf{k} v_{\mathbf{k}'}^* \delta_{\mathbf{k}\mathbf{k}'} = \delta_{\mathbf{k}\mathbf{k}'}$$

$$\{\gamma_{\mathbf{k}\uparrow}^\dagger, \gamma_{-\mathbf{k}'\downarrow}^\dagger\} = \{u_\mathbf{k} c_{\mathbf{k}\uparrow}^\dagger + v_\mathbf{k}^* c_{-\mathbf{k}\downarrow}, -v_{\mathbf{k}'}^* c_{\mathbf{k}'\uparrow} + u_{\mathbf{k}'} c_{-\mathbf{k}'\downarrow}^\dagger\}$$
$$= -u_\mathbf{k} v_{\mathbf{k}'}^* \{c_{\mathbf{k}\uparrow}^\dagger, c_{\mathbf{k}'\uparrow}\} + u_\mathbf{k} u_{\mathbf{k}'} \cancel{\{c_{\mathbf{k}\uparrow}^\dagger, c_{-\mathbf{k}'\downarrow}^\dagger\}}$$
$$- v_\mathbf{k} v_{\mathbf{k}'} \cancel{\{c_{-\mathbf{k}\downarrow}, c_{\mathbf{k}'\uparrow}\}} + v_\mathbf{k}^* u_{\mathbf{k}'} \{c_{-\mathbf{k}\downarrow}, c_{-\mathbf{k}'\downarrow}^\dagger\}$$
$$= -u_\mathbf{k} v_{\mathbf{k}'}^* \delta_{\mathbf{k}\mathbf{k}'} + u_\mathbf{k} v_{\mathbf{k}'}^* \delta_{\mathbf{k}\mathbf{k}'} = 0$$

$$\{\gamma_{\mathbf{k}\uparrow}, \gamma_{-\mathbf{k}'\downarrow}\} = \{u_\mathbf{k}^* c_{\mathbf{k}\uparrow} + v_\mathbf{k} c_{-\mathbf{k}\downarrow}^\dagger, -v_{\mathbf{k}'} c_{\mathbf{k}'\uparrow}^\dagger + u_{\mathbf{k}'} c_{-\mathbf{k}'\downarrow}\}$$
$$= -u_\mathbf{k}^* v_{\mathbf{k}'} \{c_{\mathbf{k}\uparrow}, c_{\mathbf{k}'\uparrow}^\dagger\} + u_\mathbf{k}^* u_{\mathbf{k}'} \cancel{\{c_{\mathbf{k}\uparrow}, c_{-\mathbf{k}'\downarrow}\}}$$
$$- v_\mathbf{k}^* v_{\mathbf{k}'} \cancel{\{c_{-\mathbf{k}\downarrow}^\dagger, c_{\mathbf{k}'\uparrow}^\dagger\}} + u_{\mathbf{k}'} v_\mathbf{k} \{c_{-\mathbf{k}\downarrow}^\dagger, c_{-\mathbf{k}'\downarrow}\}$$
$$= -u_\mathbf{k} v_{\mathbf{k}'}^* \delta_{\mathbf{k}\mathbf{k}'} + u_\mathbf{k} v_{\mathbf{k}'} \delta_{\mathbf{k}\mathbf{k}'} = 0$$

$$\{\gamma_{\mathbf{k}\uparrow}, \gamma_{\mathbf{k}'\uparrow}\} = \{u_\mathbf{k}^* c_{\mathbf{k}\uparrow} + v_\mathbf{k} c_{-\mathbf{k}\downarrow}^\dagger, u_{\mathbf{k}'}^* c_{\mathbf{k}'\uparrow} + v_{\mathbf{k}'} c_{-\mathbf{k}'\downarrow}^\dagger\}$$
$$= u_\mathbf{k}^* u_{\mathbf{k}'}^* \cancel{\{c_{\mathbf{k}\uparrow}, c_{\mathbf{k}'\uparrow}\}} + u_\mathbf{k}^* v_{\mathbf{k}'} \cancel{\{c_{\mathbf{k}\uparrow}, c_{-\mathbf{k}'\downarrow}^\dagger\}} + u_{\mathbf{k}'}^* v_\mathbf{k} \cancel{\{c_{-\mathbf{k}\downarrow}^\dagger, c_{\mathbf{k}'\uparrow}\}} + v_\mathbf{k} v_{\mathbf{k}'} \cancel{\{c_{-\mathbf{k}\downarrow}^\dagger, c_{-\mathbf{k}'\downarrow}^\dagger\}} = 0,$$

thus confirming the bogoliubon's fermionic character.

Then, the Hamiltonian, in terms of the bogoliubons, reads

(156)

$$\mathbf{H}_{\text{BCS}}^{\text{MFT}} = \sum_\mathbf{k} \mathbf{A}_\mathbf{k}^\dagger \mathbf{H}_k \mathbf{A}_\mathbf{k} + \text{tr}\{\mathbf{H}\} = \sum_\mathbf{k} \mathbf{B}_\mathbf{k}^\dagger \mathbf{U}_\mathbf{k}^\dagger \mathbf{H}_k \mathbf{U}_\mathbf{k} \mathbf{B}_\mathbf{k} + \text{tr}\{\mathbf{H}\}$$

$$= E_\mathbf{k}\left(\gamma_{\mathbf{k}\uparrow}^\dagger \gamma_{\mathbf{k}\uparrow} + \gamma_{\mathbf{k}\downarrow}^\dagger \gamma_{\mathbf{k}\downarrow}\right) + c$$

From the new Hamiltonian and its eigenvalues, it is clear that there are no fermion excitations possible with energy less than $|\Delta|$. The mean field parameter provides an energy gap, which is called the superconducting gap, which has a number of important consequences. The self-consistent solution is found from the gap equation, given by equation (150), by calculating $\Delta_{\mathbf{k}}$'s expectation value using the diagonalized Hamiltonian, which is equivalent to minimizing the free energy with respect to the mean field parameter $\langle \mathbf{b_k} \rangle$; thus yielding

$$
\begin{aligned}
\Delta_{\mathbf{k}} &= -\sum_{\mathbf{k}} \mathfrak{V}_{\mathbf{kk'}} \langle \mathbf{c}_{-\mathbf{k'}\downarrow}\mathbf{c}_{\mathbf{k'}\uparrow} \rangle = -\sum_{\mathbf{k}} \mathfrak{V}_{\mathbf{kk'}} \langle \left(v_{\mathbf{k'}}\gamma_{\mathbf{k'}\uparrow}^{\dagger} + u_{\mathbf{k'}}\gamma_{-\mathbf{k'}\downarrow} \right)\left(u_{\mathbf{k'}}\gamma_{\mathbf{k'}\uparrow} - v_{\mathbf{k'}}\gamma_{-\mathbf{k'}\downarrow}^{\dagger} \right) \rangle \\
&= -\sum_{\mathbf{k}} \mathfrak{V}_{\mathbf{kk'}} \left(v_{\mathbf{k'}}u_{\mathbf{k'}}\langle \gamma_{\mathbf{k'}\uparrow}^{\dagger}\gamma_{\mathbf{k'}\uparrow} \rangle - v_{\mathbf{k'}}u_{\mathbf{k'}}\langle \gamma_{-\mathbf{k'}\downarrow}\gamma_{-\mathbf{k'}\downarrow}^{\dagger} \rangle \right) \\
&= -\sum_{\mathbf{k}} \mathfrak{V}_{\mathbf{kk'}} v_{\mathbf{k'}}u_{\mathbf{k'}} \left(1 - n_{\mathfrak{FD}}(E'_{\mathbf{k}}) \right),
\end{aligned}
$$
(157)

where, in the last step, the bogoliubons' fermionic nature was used. Since these are free fermionic particles, they obey Fermi-Dirac statistics. Now, given that $\mathfrak{V}_{\mathbf{kk'}}$ is finitely-non zero only for $\varepsilon_{\mathbf{k}}, \varepsilon_{\mathbf{k}} \in [-\omega_D, \omega_D]$, and given that $\omega_D << E_F$, such taht the density of states is constant in the energy interval in question, thus yielding

$$
(158) \qquad |\Delta| = V_0 \frac{d(E_F)}{2} \int_{-\omega_D}^{\omega_D} d\varepsilon \frac{|\Delta|}{2E}\left(1 - n_{\mathfrak{FD}}(E'_{\mathbf{k}}) \right) \Rightarrow \frac{2}{V_0 d(E_F)} = \int_0^{\omega_D} d\varepsilon \frac{\tanh\left(\frac{\beta}{2}\sqrt{\varepsilon^2 + |\Delta|^2}\right)}{\sqrt{\varepsilon^2 + |\Delta|^2}},
$$

which can be numerically solved. The critical temperature is found by setting $\Delta = 0$, yielding an approximate formula

$$
kT_c = 1.13\omega_D e^{-\frac{2}{V_0}d(E_F)} \text{ and } \Delta_0 \approx 2\omega_D e^{-\frac{2}{V_0}d(E_F)}.
$$

Note that for metals, $V_0 d(E_F)$ is a very small number. Combining both equations yield the BCS prediction

$$
\frac{2\Delta_0}{kT_c} = 3.53,
$$

which is very close to the experimental result. Note that both the gap and the critical temperature are reduced by an exponential factor, compared to the bare energy scale of the interaction ω_D. This strong renormalization cannnot be obtained via pertubation theory. Moreover, no phase transition can be obtained from a perturbative approach alone, since eg. functions like $\exp(1/x)$ have no Taylor expansion at $x = 0$, ruling out any analytic connection between different phases.

3.2. **Coherent States.** In field, the basic ingredient are fields, operators $\phi(\mathbf{x})$ defined at each point in space. Thus, it is necessary to work in a basis that diagonalizes field operators, thus the coherent states arise. Moreover, coherent states comprise states with varying particle numbers, providing an extremely useful (non-orthonormal) basis for the Fock basis.

Bosonic coherent states. For a bosonic state with annihilation operator $\mathbf{b}$, the coherent states are defined as $\mathbf{b}$'s eigenstates, namely

$$
\mathbf{b}\left|z\right\rangle = z\left|z\right\rangle, \text{ where } z = |z|e^{i\phi}.
$$

Every state in the Hilbert space can thus be written as

$$
\left|z\right\rangle = \mathcal{B}(\mathbf{b}^{\dagger})\left|0\right\rangle, \quad \begin{matrix} \text{where the function } \mathcal{B}(\mathbf{b}^{\dagger}) \\ \text{has to satisfy} \end{matrix} \quad [\mathbf{b}, \mathcal{B}(\mathbf{b}^{\dagger})] = z\mathcal{B}(\mathbf{b}^{\dagger})
$$

$$
\text{Given that } [\mathbf{b}, \mathcal{B}(\mathbf{b}^{\dagger})] = \frac{\partial \mathcal{B}(\mathbf{b}^{\dagger})}{\partial \mathbf{b}^{\dagger}} \Rightarrow \frac{\partial \mathcal{B}(\mathbf{b}^{\dagger})}{\partial \mathbf{b}^{\dagger}} = z\mathcal{B}(\mathbf{b}^{\dagger}) \Rightarrow \mathcal{B}(\mathbf{b}^{\dagger}) = Ce^{z\mathbf{b}^{\dagger}}.
$$

Thus, every $\mathbf{b}$-eigenstate may be written as

$$\langle z|z\rangle = |C|^2 \langle 0| e^{z^*\mathbf{b}} e^{z\mathbf{b}^\dagger} |0\rangle \qquad\qquad \Rightarrow |z\rangle = e^{-\frac{|z|^2}{2}} e^{z\mathbf{b}^\dagger} |0\rangle$$

$$|z\rangle = Ce^{z\mathbf{b}^\dagger} |0\rangle, \text{ with} \qquad = |C|^2 e^{z^*\mathbf{b}+z\mathbf{b}^\dagger+\frac{1}{2}[z^*\mathbf{b},z\mathbf{b}^\dagger]}$$

$$= |C|^2 e^{z^* z} = 1 \Rightarrow |C|^2 = e^{-z^* z} \qquad\qquad = e^{-\frac{|z|^2}{2}} \sum_{n=0}^{\infty} \frac{z^n}{n!} |n\rangle.$$

For an N-boson system, the most general coherent state may be written as

(159)
$$|\mathbf{z}\rangle = \prod_{i=1}^{N} e^{-\frac{|z_i|^2}{2}} e^{z_i \mathbf{b}^\dagger} |0\rangle = e^{-i\sum_{i=1}^{N} \frac{|z_i|^2}{2}} e^{\sum_{i=1}^{N} z_i \mathbf{b}_i^\dagger} |0\rangle.$$

The bosonic coherent states have some interesting consequences, namely
- The coherent states are not pair-wise orthogonal to each other:

(160)
$$\langle z_1|z_2\rangle = e^{-\frac{|z_1|^2+|z_2|^2}{2}} \langle 0| e^{z_1^*\mathbf{b}} e^{z_2^*\mathbf{b}^\dagger} |0\rangle$$

$$= e^{-\frac{|z_1|^2+|z_2|^2}{2}} e^{z_1^* z_2}$$

- The coherent states form an over complete basis

$$\mathbb{1} = \int_{\mathbb{C}} \frac{dz^* dz}{\pi} |z\rangle \langle z|.$$

The trace of an operator acting on the bosonic coherent states can be written as

$$\text{Tr}\{\mathbf{A}\} = \sum_n \langle n| \mathbf{A} |n\rangle = \int \frac{dz^* dz}{\pi} |z\rangle \langle z| e^{-|z|^2} \langle z| \mathbf{A} |\mathbf{z}\rangle.$$

- The action of $\mathbf{b}^\dagger$ on a coherent state $|z\rangle$ reads

$$\mathbf{b}^\dagger |z\rangle = \mathbf{b}^\dagger \sum_{n=0}^{\infty} \frac{z^n}{n!} (\mathbf{b}^\dagger)^n |0\rangle = \sum_{n=0}^{\infty} \frac{z^n}{n!} (\mathbf{b}^\dagger)^{n+1} |0\rangle.$$

(161)
$$= \sum_{n=0}^{\infty} \frac{z^n}{(n+1)!} (\mathbf{b}^\dagger)^{(n+1)} |0\rangle = \sum_{n=1}^{\infty} \frac{z^{n-1}}{(n-1)!} (\mathbf{b}^\dagger)^{(n)} |0\rangle$$

$$= \frac{\partial}{\partial z} |z\rangle.$$

- Consider normal-ordered operator's matrix elements. Note that

$$\langle z_1| \mathbf{b}^{\dagger n} \mathbf{b}^m |z_2\rangle = z_1^{*n} z_2^m \langle z_1|z_2\rangle = z_1^{*n} z_2^m e^{z_1^* z_2},$$

which allows the matrix element of an operator to be written as

$$\mathbf{A} = \sum_{mn} a_{mn} \mathbf{b}^{\dagger n} \mathbf{b}^m \text{ such that } \langle z_1| \sum_{mn} a_{mn} \mathbf{b}^{\dagger n} \mathbf{b}^m |z_2\rangle = e^{z_1^* z_2} \sum_{mn} a_{mn} z_1^{*m} z_2^n.$$

- Any Fock state may be expanded in terms of coherent states, as

$$|\psi\rangle = \int \frac{dz^* dz}{\pi} |z\rangle \langle z|\psi\rangle = \int \frac{dz^* dz}{\pi} \phi(z^*) |z\rangle.$$

- Action of $\mathbf{b}^\dagger$ and $\mathbf{b}$ in coherent state representation reads

$$\langle z| \mathbf{b}^\dagger |z\rangle = z^* \psi(z^*) \quad \Rightarrow \mathbf{b}^\dagger \to z^*$$
$$\langle z| \mathbf{b} |z\rangle = \tfrac{\partial}{\partial z^*} \psi(z^*) \quad \Rightarrow \mathbf{b} \to \tfrac{\partial}{\partial z^*}.$$

Furthermore, the coherent state variables satisfy the bosonic commutators

$$[z_i^*, z_j^*] = \left[\frac{\partial}{\partial z_i^*}, \frac{\partial}{\partial z_j^*}\right] = 0 \text{ and } \left[\frac{\partial}{\partial z_i^*}, z_j^*\right] = \delta_{ij}.$$

- Consider the coherent-state representation of the Schrödinger equation, which reads

$$\mathbf{H}(\mathbf{b}^\dagger, \mathbf{b}) \left| \psi \right\rangle = E \left| \psi \right\rangle.$$

Then, acting from the left by $\langle z |$ will yield

$$\mathbf{H}\left(z^*, \frac{\partial}{\partial z^*} \right) \phi(z^*) = E\phi(z^*),$$

which for a typical many-body Hamiltonian reads

$$\sum_{ij} \varepsilon_{ij} z_i^* \frac{\partial}{\partial z_j^*} + \frac{1}{2} \sum_{ijkl} \langle ij | V | kl \rangle \, z_i^* z_j^* \frac{\partial}{\partial z_i^*} \frac{\partial}{\partial z_j^*}.$$

Fermionic coherent states. Fermionic coherent states may be defined in a similar way to bosonic ones. However, when considering fermionic systems, the Pauli exclusion principle must be taken into account in this formulation. The eigenstates of a fermionic annihilation should satisfy the equation

$$\mathbf{f} \left| \zeta \right\rangle = \zeta \left| \zeta \right\rangle.$$

Since, for fermions there are only two possible state occupations, the most general state may be written as $\left| \zeta \right\rangle = \alpha \left| 0 \right\rangle + \beta \left| 1 \right\rangle$, which yields

$$\mathbf{f} \left| \zeta \right\rangle = \beta \left| 1 \right\rangle = \zeta(\alpha \left| 0 \right\rangle + \beta \left| 1 \right\rangle) \Rightarrow \left| \zeta \right\rangle = \left| 0 \right\rangle \text{ with eigenvalue } 0.$$

Therefore, to obtain nontrivial fermionic coherent states, the anticommutation of annihilation operators must be considered. For two operators $\mathbf{f}_i, \mathbf{f}_j$ with coherent eigenstates

$$\mathbf{f}_i \left| \boldsymbol{\zeta} \right\rangle = \zeta_i \left| \boldsymbol{\zeta} \right\rangle \text{ and } \mathbf{f}_j \left| \boldsymbol{\zeta} \right\rangle = \zeta_j \left| \boldsymbol{\zeta} \right\rangle \Rightarrow \text{ if } \{\mathbf{f}_i, \mathbf{f}_j\} = 0 \Leftrightarrow \{\zeta_i, \zeta_j\} = 0.$$

This equation implies that $\zeta_i^2 = 0$ as well. This implies that the ζ_i-numbers are no ordinary numbers but are the elements of a Grassmann algebra, and such objects are identified as Grassmann numbers.

Grassmann Algebras

Formally, let $\mathbf{V}$ be an n-dimensional complex vector space with basis $\{\theta_i\}_{i=1}^n$. The Grassmann algebra, whose Grassman variables are $\{\theta_i\}_{i=1}^n$, is defined as the exterior algebra of $\mathbf{V}$, namely

(162)
$$\begin{aligned}
\Lambda(\mathbf{V}) &= \mathbb{C} \oplus \mathbf{V} \oplus (\mathbf{V} \wedge \mathbf{V}) \oplus (\mathbf{V} \wedge \mathbf{V} \wedge \mathbf{V}) \oplus \cdots \oplus (\textstyle\bigwedge^n \mathbf{V}) \\
&= \mathbb{C} \oplus \bigoplus k = 1^n \textstyle\bigwedge^k \mathbf{V},
\end{aligned}$$

where $\wedge$ is the exterior product. The individual elements of this algebra are the Grassmann numbers, which in their most general form may be written as

(163)
$$\begin{aligned}
z &= c_0 + \sum_{k=1}^n \sum_{\mathrm{i}_k^\uparrow} c_{i_1, i_2, \cdots, i_k} \theta_{i_1} \theta_{i_2} \cdots \theta_{i_k}, \\
&= c_0 + \sum_i c_i \theta_i + \sum_{i<j} c_i c_j \theta_i \theta_j + \sum_{i<j<k} c_i c_j c_k \theta_i \theta_j \theta_k + \cdots
\end{aligned}$$

where $\mathrm{i}^\uparrow = (i_1, i_2, \cdots, i_k)$ are strictly increasing k-tuples, with $1 \le i \le n, 1 \le j \le k$, and where $c_{i_1, i_2, \cdots, i_k}$ are complex, completely antisymmetric tensors of rank k. Again, the θ_i, the $\theta_i \wedge \theta_j = \theta_i \theta_j$ (subject to $i < j$) and larger finite products, can be seen playing the role of a basis vector of Λ-subspaces. The Grassmann algebra generated by n linearly independent Grassmann variables has dimension 2^n, since the $(n+1)$-fold product of variables must vanish. The dimension of $\bigwedge^k \mathbf{V}$ is given by the n choose k binomial coefficient. In the special case of a infinite dimensional $\mathbf{V}$-space, the previous series does not terminate, thus yielding a $\bigwedge^\infty \mathbf{V}$.

For Grassmann algebras with an even number of generators, $n = 2p$, a conjugation can be defined by selecting p generators θ_i and associating $\theta^\star$ to each of them, such that

$$(\theta_i)^\star = \theta_i^\star \qquad\qquad\qquad (\theta_i^\star)^\star = \theta_i.$$

In a fermionic field theory, the Grassmann algebra is defined by the set of its generating elements $\{1, \zeta_1, \zeta_1^\star, \cdots, \zeta_n, \zeta_n^\star\}$, which satisfy

(164)
$$\{\zeta_i, \zeta_j\} = \{\zeta_j^\star, \zeta_j^\star\} = \{\zeta_i, \zeta_j^\star\} = \zeta_i^2 = \zeta_i^{\star 2} = 0,$$
$$\{\zeta_i, \mathbf{f}_i\} = \{\zeta_i^\star, \mathbf{f}_j\} = \{\zeta_i, \mathbf{f}_j^\dagger\} = \{\zeta_i^\star, \mathbf{f}_j^\dagger\} = 0.$$

Then, the basis set of the Grassmann algebra consists of all distinct products of its generators, namely $\{1, \zeta_1, \cdots, \zeta_1 \zeta_2, \cdots \zeta_n\}$.

For a two-dimensional Grassmann algebra, with generators $\{\zeta, \zeta^\star\}$, all analytic functions of a Grassmann variables reduce to a first degree polynomial $\phi \in \mathbb{C}[x]$, since $\zeta^2 = 0$,

$$\psi(\zeta) = \psi_0 + \psi_1 \zeta \Leftrightarrow \phi^\star(\zeta) = \phi_0{}^\star + \phi_1^\star \zeta^\star.$$

The most general bilinear function, acting on both ζ and $\zeta^\star$, is given by

$$F(\zeta, \zeta^\star) = f_0 + f_1 \zeta + f_1^\star \zeta^\star + f_{12} \zeta^\star \zeta.$$

Grassmann calculus

- Differentiation of Grassmann variables:

  Since analytic functions of Grassmann variables have such a simple structure, differentiation is just as simple and is given by the usual complex derivative. In effect, let $\psi(\zeta) = \psi_0 + \psi_1 \zeta$. Then, the derivatives are defined to be

$$\partial_\zeta \psi(\zeta) = \psi_1, \qquad\qquad \partial_{\zeta^\star} \psi^\star(\zeta) = \phi_1^\star, \qquad\qquad \partial_\zeta(\zeta^\star \zeta) = -\bar{\zeta},$$

  where in the last equality, $\partial_{\zeta^\star}(\zeta^\star \zeta) = -\partial_{\zeta^\star}(\zeta \zeta^\star) = -\zeta^\star$. Now let $F(\zeta, \zeta^\star) = f_0 + f_1 \zeta + f_1^\star \zeta^\star + f_{12} \zeta^\star \zeta$, then

$$\partial_\zeta F(\zeta, \zeta^\star) = f_1 - f_{12} \zeta^\star, \qquad \partial_{\zeta^\star} F(\zeta, \zeta^\star) = f_1 + f_{12} \zeta, \qquad \partial_{\zeta^\star} \partial_\zeta F(\zeta, \zeta^\star) = -f_{12} = -\partial_\zeta \partial_{\zeta^\star} F(\zeta, \zeta^\star),$$

  which in turn implies that $\{\partial_{\zeta^\star}, \partial_\zeta\} = 0$. The chain rule for differentiation then reads

$$\partial_\zeta F(\psi) = \partial_\zeta \psi \, \partial_\psi F,$$

  wherein, contrary to ordinary differentiation, the order of the terms matters.

- Integration over Grassmann variables: the Berezin integral.
  Integration and differentiation are identical for Grassmann variables, wherein

$$\partial_{\zeta_i} F(\boldsymbol{\zeta}) = \int d\zeta_i \, F(\boldsymbol{\zeta}).$$

  This definition thus ensures that two most fundamental of properties of ordinary integrals over functions vanishing at infinity are satisfied

  – The integral of an exact differential form is zero:

$$\int d\zeta_i \, \partial_{\zeta_i} F(\boldsymbol{\zeta}) = 0.$$

– The integral over ζ_i of $F(\boldsymbol{\zeta})$ does not depend on ζ_i, thus having a vanishing ∂_{ζ_i}-derivative,

$$\partial_{\zeta_i} \int d\zeta_i \, F(\boldsymbol{\zeta}) = 0.$$

Both properties follow from the Grassmann derivative and its nilpotence. Another immediate consequence is that

$$\int d\zeta \, 1 = 0.$$

Thus, for a Grassmann algebra with a single generator,

(165)
$$F(\zeta) = c_0 + c_1\zeta \Rightarrow \int d\zeta \, F(\zeta) = c_1 = \partial_\zeta \, F(\zeta).$$

Properties of Fermionic coherent states

A fermionic coherent state is then defined as

(166)
$$|\zeta\rangle = \mathcal{N} e^{-\sum_\alpha \zeta_\alpha \mathbf{f}_\alpha^\dagger} |0\rangle = \mathcal{N} \prod_\alpha (1 - \zeta_\alpha \mathbf{f}_\alpha^\dagger) |0\rangle$$

(167)
$$\langle\zeta| = \mathcal{N} \langle 0| e^{-\sum_\alpha \zeta_\alpha^\star \mathbf{f}_\alpha} = \mathcal{N} \prod_\alpha \langle 0| (1 - \zeta_\alpha^\star \mathbf{f}_\alpha, \text{ since}$$

(168)
$$\mathbf{f}_\alpha |\zeta\rangle = \mathcal{N} \mathbf{f}_\alpha \prod_\beta (1 - \zeta_\beta \mathbf{f}_\beta^\dagger) |0\rangle$$
$$= \mathcal{N} \prod_{\beta\neq\alpha} (1 - \zeta_\beta \mathbf{f}_\beta^\dagger) \mathbf{f}_\alpha (1 - \zeta_\alpha \mathbf{f}_\alpha^\dagger) |0\rangle$$
$$= \mathcal{N} \prod_{\beta\neq\alpha} (1 - \zeta_\beta \mathbf{f}_\beta^\dagger) \zeta_\alpha (1 - \zeta_\alpha \mathbf{f}_\alpha^\dagger) |0\rangle$$
$$= \mathcal{N} \zeta_\alpha \prod_\beta (1 - \zeta_\beta \mathbf{f}_\beta^\dagger) |0\rangle$$
$$= \zeta_\alpha |\zeta\rangle \,,$$

$$\mathbf{f}_\alpha^\dagger |\zeta\rangle = \mathcal{N} \mathbf{f}_\alpha^\dagger (1 - \zeta_\alpha \mathbf{f}_\alpha^\dagger) \prod_{\beta\neq\alpha} (1 - \zeta_\beta \mathbf{f}_\beta^\dagger) |0\rangle$$
$$= \mathcal{N} \mathbf{f}_\alpha^\dagger \prod_{\beta\neq\alpha} (1 - \zeta_\beta \mathbf{f}_\beta^\dagger) |0\rangle$$
$$= -\mathcal{N} \frac{\partial}{\partial\zeta_\alpha} (1 - \zeta_\alpha \mathbf{f}_\alpha^\dagger) \prod_{\beta\neq\alpha} (1 - \zeta_\beta \mathbf{f}_\beta^\dagger) |0\rangle$$
$$= -\frac{\partial}{\partial\zeta_\alpha} |\zeta\rangle \,.$$

(169)
$$\Rightarrow \quad \begin{array}{ll} \mathbf{f}_\alpha |\zeta\rangle = \zeta_\alpha |\zeta\rangle & \mathbf{f}_\alpha^\dagger |\zeta\rangle = -\partial_{\zeta_\alpha} |\zeta\rangle \\[6pt] \langle\zeta| \mathbf{f}_\alpha^\dagger = \langle\zeta| \zeta_\alpha^\star & \langle\zeta| \mathbf{f}_\alpha = +\partial_{\zeta_\alpha^\star} |\zeta\rangle \end{array}$$

For a Grassmann-variable, η, the associated inner product is defined as

(170)
$$\langle\zeta|\eta\rangle = \langle 0| e^{-\zeta^\star \mathbf{f}} e^{\eta \mathbf{f}^\dagger} |0\rangle = 1 + \zeta^\star \eta = e^{\zeta^\star \eta}.$$

In particular, if $\eta = \zeta$, the normalization factor is $\mathcal{N} = e^{-\frac{\zeta^\star \zeta}{2}}$. Then, generalizing to N fermions, the coherent states read

$$|\boldsymbol{\zeta}\rangle = \prod_{i=1}^N e^{-\frac{\zeta_i^\star \zeta_i}{2}} e^{-\zeta_i \mathbf{f}_i^\dagger} |0\rangle = e^{-\sum_i \frac{\zeta_i^\star \zeta_i}{2}} e^{-\sum_i \zeta_i \mathbf{f}_i^\dagger} |0\rangle \,, \text{ since } [\zeta_i \mathbf{f}_i^\dagger, \zeta_j \mathbf{f}_j^\dagger] = 0.$$

Fermions obey a closure relation as well.

(171)
$$\int \prod_{i=1}^N d\zeta_i^\star d\zeta_i \, |\boldsymbol{\zeta}\rangle \langle\boldsymbol{\zeta}| = \mathbb{1}_{\mathcal{F}}.$$

Fermionic closure relation

In effect, said relation holds if and only if

$$\langle \alpha_1, \cdots, \alpha_n | \int \prod_{i=1}^{N} d\zeta_i^\star d\zeta_i \ |\boldsymbol{\zeta}\rangle \langle \boldsymbol{\zeta}|\beta_1, \cdots, \beta_m\rangle = \langle \alpha_1, \cdots, \alpha_n|\beta_1, \cdots, \beta_m\rangle .$$

Said overlap is

$$\langle \alpha_1, \cdots, \alpha_n|\beta_1, \cdots, \beta_m\rangle = \delta_{nm}(-1)^p \delta_{\alpha_{i_1}\beta_1} \cdots \delta_{\alpha_{i_n}\beta_m}, \quad \begin{array}{l} \text{where } p \text{ is the parity} \\ \text{of the permutation} \end{array} P \left(\begin{array}{ccc} \alpha_1, & \cdots, & \alpha_n \\ \alpha_{i_i}, & \cdots, & \alpha_{i_n} \end{array} \right).$$

Note that $\langle \alpha_1, \cdots, \alpha_n|\zeta\rangle = \langle 0| \mathbf{f}_{\alpha_n} \cdots \mathbf{f}_{\alpha_1} |\zeta\rangle = \zeta_{\alpha_n} \cdots \zeta_{\alpha_1}$, thus leading to the following matrix element

(172)

$$\langle \alpha_1, \cdots, \alpha_n | \int \prod_{i=1}^{N} d\zeta_i^\star d\zeta_i \ |\boldsymbol{\zeta}\rangle \langle \boldsymbol{\zeta}|\beta_1, \cdots, \beta_m\rangle = \int \prod_{i=1}^{N} d\zeta_i^\star d\zeta_i \ \langle \alpha_1, \cdots, \alpha_n|\boldsymbol{\zeta}\rangle \langle \boldsymbol{\zeta}|\beta_1, \cdots, \beta_m\rangle$$

$$= \int \prod_{i=1}^{N} d\zeta_i^\star d\zeta_i \ \prod_{j=1}^{n} \langle \alpha_j|\boldsymbol{\zeta}\rangle \prod_{k=1}^{m} \langle \boldsymbol{\zeta}|\beta_k\rangle$$

$$= \int \prod_{i=1}^{N} d\zeta_i^\star d\zeta_i \ \prod_{j=1}^{n} \left(\langle 0| \mathbf{f}_{\alpha_j} |\zeta\rangle \right) \prod_{k=1}^{m} \left(\langle \zeta| \mathbf{f}_{\beta_k} |0\rangle \right)$$

$$= \int \prod_{i=1}^{N} d\zeta_i^\star d\zeta_i \ \prod_{j=1}^{n} e^{-\sum_\ell \frac{\zeta_\ell^\star \zeta_\ell}{2}} \left(\zeta_{\alpha_j} \langle 0|\zeta\rangle \right) \prod_{k=1}^{m} e^{-\sum_{\ell'} \frac{\zeta_{\ell'}^\star \zeta_{\ell'}}{2}} \left(\zeta_{\beta_k^\star} \langle \zeta|0\rangle \right)$$

$$= \int \prod_{i=1}^{N} d\zeta_i^\star d\zeta_i \ e^{-\sum_\ell \zeta_\ell^\star \zeta_\ell} \prod_{j=1}^{n} \zeta_{\alpha_j} \prod_{k=1}^{m} \zeta_{\beta_k^\star}$$

$$= \int \prod_{i=1}^{N} d\zeta_i^\star d\zeta_i \ \sum_\ell \left(1 - \zeta_\ell^\star \zeta_\ell \right) \prod_{j=1}^{n} \zeta_{\alpha_j} \prod_{k=1}^{m} \zeta_{\beta_k^\star}$$

with the following integrals arising for a particular $|\gamma\rangle$-state

(173)

$$\int d\zeta_\gamma^\star d\zeta_\gamma \left(1 - \zeta_\ell^\star \zeta_\ell \right) \left\{ \begin{array}{c} \zeta_\gamma^\star \zeta_\gamma \\ \zeta_\gamma^\star \\ \zeta_\gamma \\ 1 \end{array} \right\} = \left\{ \begin{array}{c} 1 \\ 0 \\ 0 \\ 1 \end{array} \right\}.$$

Therefore, this integral is non-vanishing only if each $|\gamma\rangle$-state is occupied in both $\langle \alpha_1, \cdots, \alpha_n|$ and $\langle \beta_1, \cdots, \beta_m|$ or unoccupied in both states. In particular, if $m = n$, $\boldsymbol{\alpha}$ is a permutation of $\boldsymbol{\beta}$.

Overcompleteness allows to define Grassmann coherent state representation for Fock states. A Fock-state $|\psi\rangle$ can be expanded in terms of fermionic coherent states as

$$|\psi\rangle = \int \prod_{i=1}^{N} d\zeta_i^\star d\zeta_i \ \psi(\boldsymbol{\zeta})^\star |\boldsymbol{\zeta}\rangle, \text{ where } \psi(\boldsymbol{\zeta})^\star = e^{-\sum_\alpha \zeta_\alpha^\star \zeta_\alpha} \langle \zeta|\psi\rangle$$

wherein, the following matrix elements readily are found to be

$$\langle\boldsymbol{\zeta}|\,\mathbf{f}_i\,|\psi\rangle = \int \prod_{i=1}^{N} d\zeta_i^\star d\zeta_i \,\langle\boldsymbol{\zeta}|\,\psi(\boldsymbol{\zeta})^\star\mathbf{f}_i\,|\boldsymbol{\zeta}\rangle \qquad \langle\boldsymbol{\zeta}|\,\mathbf{f}_i^\dagger\,|\psi\rangle = \int \prod_{i=1}^{N} d\zeta_i^\star d\zeta_i \,\langle\boldsymbol{\zeta}|\,\psi(\boldsymbol{\zeta})^\star\mathbf{f}_i^\dagger\,|\boldsymbol{\zeta}\rangle$$

$$(174)\qquad = \int \prod_{i=1}^{N} d\zeta_i^\star d\zeta_i \,\partial_{\zeta_i^\star}\psi(\boldsymbol{\zeta})^\star\,\langle\boldsymbol{\zeta}|\boldsymbol{\zeta}\rangle \qquad\qquad = \int \prod_{i=1}^{N} d\zeta_i^\star d\zeta_i \,\zeta_i^\star\psi(\boldsymbol{\zeta})^\star\,\langle\boldsymbol{\zeta}|\boldsymbol{\zeta}\rangle$$

$$= \partial_{\zeta^\star}\psi(\boldsymbol{\zeta})^\star \qquad\qquad\qquad\qquad = \zeta_i^\star\psi(\boldsymbol{\zeta})^\star$$

The matrix elements of a normal-ordered operator $\mathbf{A}(\mathbf{f}_i^\dagger, \mathbf{f}_i)$ is

$$\langle\boldsymbol{\zeta}|\,\mathbf{A}(\{\mathbf{f}_i^\dagger\}, \{\mathbf{f}_i\})\,|\boldsymbol{\zeta}\rangle = e^{-\frac{|\zeta|^2 + |\zeta'|^2}{2}} e^{\sum_i \zeta_i^\star \zeta_i'} \mathbf{A}(\{\zeta_i^\star\}, \{\zeta_i\}).$$

In the case of the fermionic number operator, $\mathbf{N} = \sum_i \mathbf{f}_i^\dagger\mathbf{f}_i$ gives its expectation value

$$\frac{\langle\boldsymbol{\zeta}|\,\mathbf{N}\,|\boldsymbol{\zeta}\rangle}{\langle\boldsymbol{\zeta}|\boldsymbol{\zeta}\rangle}.$$

3.3. **Gaussian Integrals.** Quantum mechanical actions, commonly encountered in many-body path integral formalism, frequently involve exponential or complex or Grassmann variables. For quadratic-form actions, these integrals are simple generalizations of Gaussian integrals.

- **Real variables**:

If the integral can be written as

$$(175)\qquad \mathcal{Z}(J) = \int_{\Omega\subseteq\mathbb{R}^{3N}} \prod_{i=1}^{N} d\mathbf{x}_i \, e^{-\frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} + \mathbf{J}^\mathsf{T}\mathbf{x}}, \quad \text{with} \quad \begin{matrix} \mathbf{A} = \mathbf{A}^\mathsf{T} \\ \sigma(\mathbf{A}) > 0 \end{matrix}$$

Thus, the $3N$-dimensional real symmetric and positive defined matrix $\mathbf{A}$ can be diagonalized via an orthonormal transformation $\mathbf{A} = \mathbf{O}^\mathsf{T} D\mathbf{O}$, with $\mathbf{O} \in O(3N)$ and $D_{ii} > 0$. Now, using these facts, the exponentiated factor can be rewritten as a bilinear form, without the linear term, by "completing the square". In effect, let $\mathbf{y} = \mathbf{O}(\mathbf{x} - \mathbf{A}^{-1}\mathbf{J})$, or equivalently $\mathbf{x} = \mathbf{O}^{-1}\mathbf{y} + \mathbf{A}^{-1}\mathbf{J}$

$$
\begin{aligned}
-\frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} + \mathbf{J}^\mathsf{T}\mathbf{x} &= -\frac{1}{2}\left(\mathbf{O}^{-1}\mathbf{y} + \mathbf{A}^{-1}\mathbf{J}\right)^\mathsf{T}\mathbf{A}\left(\mathbf{O}^{-1}\mathbf{y} + \mathbf{A}^{-1}\mathbf{J}\right) + \mathbf{J}^\mathsf{T}\left(\mathbf{O}^{-1}\mathbf{y} + \mathbf{A}^{-1}\mathbf{J}\right) \\
&= -\frac{1}{2}\left(\mathbf{y}^\mathsf{T}\mathbf{O} + \mathbf{J}^\mathsf{T}(\mathbf{A}^{-1})^\mathsf{T}\right)\mathbf{A}\left(\mathbf{O}^{-1}\mathbf{y} + \mathbf{A}^{-1}\mathbf{J}\right) + \mathbf{J}^\mathsf{T}\left(\mathbf{O}^{-1}\mathbf{y} + \mathbf{A}^{-1}\mathbf{J}\right) \\
&= -\frac{1}{2}\left[\mathbf{y}^\mathsf{T}\mathbf{O}\mathbf{A}\mathbf{O}^{-1}\mathbf{y} + \mathbf{y}^\mathsf{T}\mathbf{O}\mathbf{A}\mathbf{A}^{-1}\mathbf{J} + \mathbf{J}^\mathsf{T}(\mathbf{A}^{-1})^\mathsf{T}\mathbf{A}\mathbf{O}^{-1}\mathbf{y} + \mathbf{J}^\mathsf{T}(\mathbf{A}^{-1})^\mathsf{T}\mathbf{A}\mathbf{A}^{-1}\mathbf{J}\right] \\
&\quad + \mathbf{J}^\mathsf{T}\mathbf{O}^{-1}\mathbf{y} + \mathbf{J}^\mathsf{T}\mathbf{A}^{-1}\mathbf{J} \\
(176)\qquad &= -\frac{1}{2}\left[\mathbf{y}^\mathsf{T}D\mathbf{y} + \mathbf{y}^\mathsf{T}\mathbf{O}\mathbf{J} + \mathbf{J}^\mathsf{T}\mathbf{A}^{-1}\mathbf{A}\mathbf{O}^{-1}\mathbf{y} + \mathbf{J}^\mathsf{T}\mathbf{A}^{-1}\mathbf{J}\right] + \mathbf{J}^\mathsf{T}\mathbf{O}^{-1}\mathbf{y} + \mathbf{J}^\mathsf{T}\mathbf{A}^{-1}\mathbf{J} \\
&= -\frac{1}{2}\left[\mathbf{y}^\mathsf{T}D\mathbf{y} + (\mathbf{O}\mathbf{J})^\mathsf{T}\mathbf{y}^\mathsf{T} + \mathbf{J}^\mathsf{T}\mathbf{O}^{-1}\mathbf{y} + \mathbf{J}^\mathsf{T}\mathbf{A}^{-1}\mathbf{J}\right] + \mathbf{J}^\mathsf{T}\mathbf{O}^{-1}\mathbf{y} + \mathbf{J}^\mathsf{T}\mathbf{A}^{-1}\mathbf{J} \\
&= -\frac{1}{2}\left[\mathbf{y}^\mathsf{T}D\mathbf{y} + (\mathbf{J}^\mathsf{T}\mathbf{O}^\mathsf{T}\mathbf{y})^\mathsf{T} + \mathbf{J}^\mathsf{T}\mathbf{O}^{-1}\mathbf{y} + \mathbf{J}^\mathsf{T}\mathbf{A}^{-1}\mathbf{J}\right] + \mathbf{J}^\mathsf{T}\mathbf{O}^{-1}\mathbf{y} + \mathbf{J}^\mathsf{T}\mathbf{A}^{-1}\mathbf{J} \\
&= -\frac{1}{2}\mathbf{y}^\mathsf{T}D\mathbf{y} - \frac{1}{2}(\mathbf{J}^\mathsf{T}\mathbf{O}^\mathsf{T}\mathbf{y})^\mathsf{T} + \frac{1}{2}\mathbf{J}^\mathsf{T}\mathbf{O}^{-1}\mathbf{y} + \frac{1}{2}\mathbf{J}^\mathsf{T}\mathbf{A}^{-1}\mathbf{J} \\
&= -\frac{1}{2}\mathbf{y}^\mathsf{T}D\mathbf{y} + \frac{1}{2}\mathbf{J}^\mathsf{T}\mathbf{O}^{-1}\mathbf{y} + \frac{1}{2}\mathbf{J}^\mathsf{T}\mathbf{A}^{-1}\mathbf{J}
\end{aligned}
$$

$$(177) \qquad \Rightarrow \mathcal{Z}(J) = e^{\frac{1}{2}\mathbf{J}^{\mathsf{T}}\mathbf{A}^{-1}\mathbf{J}} \prod_{i=1}^{N} \int_{\mathbb{R}} dy \, e^{-\frac{y^2 D_{ii}}{2}} = {\frac{1}{2}\mathbf{J}^{\mathsf{T}}\mathbf{A}^{-1}\mathbf{J}} \sqrt{\frac{(2\pi)^N}{\det A}}$$

- **Complex variables**:
  Consider now

$$(178) \qquad \begin{aligned} \mathcal{Z}(\mathbf{J}^*, \mathbf{J}) &= \int_{\Omega \subseteq \mathbb{C}^{6N}} \prod_{i=1}^{N} \frac{dz_i^* dz_i}{2\pi i} \, e^{-\sum_{ij} z_i \mathbf{A}_{ij} z_j + \sum_i J_i^* z_i + \text{h.c.}}, \text{ with } \begin{array}{c} \mathbf{A} = \mathbf{A}^\dagger \\ \sigma(\mathbf{A}) > 0 \end{array} \\ &= \int_{\Omega \subseteq \mathbb{C}^{6N}} \prod_{i=1}^{N} \frac{dz_i^* dz_i}{2\pi i} \, e^{-\mathbf{z}^\dagger \mathbf{A} \mathbf{z} + (\mathbf{J}^\dagger \mathbf{z} + \text{h.c.})} \end{aligned}$$

Thus, $\mathbf{A}$ may be diagonalized via a unitary transformation, $\mathbf{D} = U^\dagger A U$, which induce the following change of variables

$$(179) \qquad \mathbf{z}' = U(\mathbf{z} - \mathbf{A}^{-1}\mathbf{J}), \qquad\qquad \mathbf{z}'^\dagger = U^\dagger(\mathbf{z}^\dagger - \mathbf{J}^\dagger \mathbf{A}^{-1}),$$

yielding

$$(180) \qquad \mathcal{Z}(\mathbf{J}^*, \mathbf{J}) = e^{\mathbf{J}^\dagger \mathbf{A}^{-1}\mathbf{J}} \int_{\mathbb{C}} \prod_{i=1}^{N} \frac{dz_i^* dz_i}{2\pi i} \, e^{-\mathbf{z}'^\dagger \mathbf{D} \mathbf{z}'} = e^{\mathbf{J}^\dagger \mathbf{A}^{-1}\mathbf{J}} (\det A)^{-1}$$

- **Grasmannian integrals**:

  A Gaussian integral involving a conjugate pair of Grassmann variables can be written as

$$\int d\zeta^\star d\zeta \, e^{-\zeta^\star a \zeta} = \int d\zeta^\star d\zeta \, (1 - \zeta^\star a \zeta) = a.$$

  By analogy, Grassmann gaussian integrals may be written as

$$(181) \qquad \mathcal{Z}(\boldsymbol{\zeta}^\star, \boldsymbol{\zeta}) = \int \prod_i d\zeta_i^\star d\zeta_i \, e^{-\boldsymbol{\zeta}^\star \mathbf{A} \boldsymbol{\zeta} + (\boldsymbol{\varepsilon}^\star \boldsymbol{\zeta} + \text{h.c.})}.$$

  To solve this Grassmann integral, the $\mathbf{A}$-matrix must be diagonalized by a suitable change of Grassmann-variables, which themselves have not a trivial transformation.

Transformation of Grassmann variables

Consider the integral[a]

$$(182) \qquad \begin{aligned} \mathcal{I} &= \int d\zeta \, F(\zeta), \text{ subject to a transformation } \zeta \to \alpha\zeta' + \beta, \alpha, \beta \in \mathbb{C}. \\ \Rightarrow \mathcal{I} &= \partial_\zeta \, F(\zeta) = \partial_\zeta(\zeta') \partial_{\zeta'} F(\zeta') = \frac{1}{\alpha}\partial_{\zeta'} F(\zeta') = \frac{1}{\alpha} \int d\zeta \, F(\alpha\zeta' + \beta). \end{aligned}$$

Therefore, the jacobian of the transformation is $\frac{1}{\alpha}$, instead of α for ordinary (bosonic) transformation. This result can then be generalized to a multidimensional integral

---

[a]Note that, if θ is a Grassmann variable, then $\theta = a\eta$ is another Grassmann variable

$$1 = \int d\theta \, \theta = \int d(a\eta) \, a\eta \Rightarrow d(a\eta) = \frac{1}{a}d\eta.$$

$$\mathcal{I} = \int \prod_i d\zeta \ F(\boldsymbol{\zeta}) = \prod_i \partial_{\zeta_i} F(\boldsymbol{\zeta})$$

$$= \prod_i \left(\sum_j (\partial_{\zeta_i} \zeta_j') \partial_{\zeta_j'} \right) F(\boldsymbol{\zeta})$$

(183)

$$\overset{\boldsymbol{\zeta}'=\mathbf{M}\boldsymbol{\zeta}}{=} \prod_i \left(\sum_j \mathbf{M}_{ji} \partial_{\zeta_j'} \right) F(\boldsymbol{\zeta})$$

Given that $\zeta_{i_1} \cdots \zeta_{i_n} = \varepsilon_{i_1 \cdots i_n} \zeta_{i_1}' \cdots \zeta_{i_n}'$, then

$$= \prod_i \left(\varepsilon^{i_1 \cdots i_n} \mathbf{M}_{1i_1} \cdots \mathbf{M}_{ni_n} \partial_{\zeta_{i_1}} \cdots \partial_{\zeta_{i_n}} \right) F(\boldsymbol{\zeta})$$

(184)

$$= \det(\partial_{\zeta_j} \zeta_i') \prod_i \partial_{\zeta_j} \zeta_i' F(\boldsymbol{\zeta})$$

$$\Rightarrow \prod_i d\zeta_i = \det(\partial_{\zeta_j} \zeta_i') \prod_i d\zeta_i',$$

the jacobian being the inverse of the determinant $\det(\partial_{\zeta_j} \zeta_i')$.

With this in mind, a wide class of Grassmannian integrals may be solved, for example,

(185)
$$\mathcal{Z}(\boldsymbol{\zeta}^\star, \boldsymbol{\zeta}) = \int \prod_i d\zeta_i^\star d\zeta_i \ e^{-\boldsymbol{\zeta}^\star \mathbf{A} \boldsymbol{\zeta}} = \det(\mathbf{A}) \int \prod_i d\zeta_i^\star d\zeta_i \ e^{\boldsymbol{\zeta}^\star \boldsymbol{\zeta}}$$

$$= \det(\mathbf{A}) \int \prod_i d\zeta_i^\star d\zeta_i \ (1 - \boldsymbol{\zeta}^\star \boldsymbol{\zeta}) = \det(\mathbf{A}),$$

or, for the more general integral given by equation (181)

(186)
$$\mathcal{Z}(\boldsymbol{\zeta}^\star, \boldsymbol{\zeta}) = \det(\mathbf{A}) e^{-\boldsymbol{\zeta}^\star \mathbf{A}^{-1} \boldsymbol{\zeta}}.$$

3.4. **Photons: coupling to electrons.** Phonons are quantized vibrations and play a fundamental role in sound physics, specific heat, elasticity and electrical resistivity of solids. More surprisingly, the electron-phonon coupling is the cause of conventional superconductivity. There are two distinct models for electron-phonon coupling, the jellium model, where the ions are represented by a smeared-out continuous positive background, and the lattice model where the ions oscillate around their equilibrium positions forming a crystal lattice.

Phonons are basically harmonic oscillators, they are bosons. Moreover, these naturally occur at finite temperature, so the thermal distribution function for bosons will be used.

The Jellium Model. Consider the Hamiltonian

(187)

$$
\mathbf{H} = \mathbf{H}_+ + \mathbf{H}_{+e} + \mathbf{H}_e, \text{ where }
\begin{aligned}
\mathbf{H}_+ &= \tfrac{1}{2} \int_{\mathbb{R}^6} d\mathbf{x} d\mathbf{x}' \, \rho_+(\mathbf{x})\rho_+(\mathbf{x}') \frac{e^{-\mu|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x}-\mathbf{x}'|} = \frac{e^2}{2}\left(\frac{N}{\Omega}\right)^2 \int_{\mathbb{R}^3} d\mathbf{x}' \int_{\mathbb{R}^3} \frac{e^{-\mu r}}{r} = \frac{4\pi e^2 N^2}{2\mu^2 \Omega} \\
\mathbf{H}_{+e} &= -e \sum_i \int_{\mathbb{R}^3} d\mathbf{x} \, \rho_+(\mathbf{x}) \frac{e^{-\mu|\mathbf{x}-\mathbf{x}_i|}}{|\mathbf{x}-\mathbf{x}_i|} = -\frac{e^2 N}{\Omega} \sum_i \int_{\mathbb{R}^3} d\mathbf{x} \frac{e^{-\mu|\mathbf{x}-\mathbf{x}_i|}}{|\mathbf{x}-\mathbf{x}_i|} = -\frac{e^2 N^2}{\Omega}\frac{4\pi}{\mu^2}
\end{aligned}
$$

and consider the electron Hamilton

(188)
$$
\mathbf{H}_e = \sum_i \frac{\mathbf{P}^2}{2m_e} + \frac{e^2}{2}\sum_{i\neq j} \frac{e^{-\mu|\mathbf{x}-\mathbf{x}_i|}}{|\mathbf{x}-\mathbf{x}_i|}, \text{ 2nd } \underset{\longrightarrow}{\text{Quant.}} \ \mathbf{T} = \sum_i \frac{\hbar^2 k_i^2}{2m_e}\mathbf{c}_i^\dagger \mathbf{c}_i,
$$

since the single -electron wavefunctions are just plane waves, with the matrix element for the two-electron repulsion given by

(189)
$$
\begin{aligned}
\left\langle \mathbf{k}_4\sigma_i, \mathbf{k}_3\sigma_j \Big| \frac{e^{-\mu|\mathbf{x}_i-\mathbf{x}_j|}}{|\mathbf{x}_i-\mathbf{x}_j|} \Big| \mathbf{k}_2\sigma_j, \mathbf{k}_1\sigma_i \right\rangle &= \frac{1}{\Omega^2}\int_{\mathbb{R}^6} d\mathbf{x}_i d\mathbf{x}_j \, e^{i(\mathbf{k}_1-\mathbf{k}_4)\cdot\mathbf{x}_i} e^{i(\mathbf{k}_2-\mathbf{k}_3)\cdot\mathbf{x}_j} \frac{e^{-\mu|\mathbf{x}_i-\mathbf{x}_j|}}{|\mathbf{x}_i-\mathbf{x}_j|} \\
&= \frac{1}{\Omega}\delta_{\mathbf{k}_1+\mathbf{k}_2-\mathbf{k}_3-\mathbf{k}_4} \int_{\mathbb{R}^3} d(\mathbf{x}_i-\mathbf{x}_j) e^{-i\mathbf{q}\cdot(\mathbf{x}_i-\mathbf{x}_j)}\frac{e^{-\mu r}}{r} \\
&= \frac{1}{\Omega}\delta_{\mathbf{k}_1+\mathbf{k}_2-\mathbf{k}_3-\mathbf{k}_4}\frac{4\pi}{\mu^2+q^2}. \\
\Rightarrow \mathbf{H}_e &= \sum_{\substack{\mathbf{k}\in\Lambda \\ \sigma\in\mathbb{C}^2}} \frac{\hbar^2 k^2}{2m_e}\mathbf{c}_{\mathbf{k}\sigma}^\dagger \mathbf{c}_{\mathbf{k}\sigma} + \sum_{\substack{\mathbf{k},\mathbf{k}',\mathbf{q}\in\Lambda \\ \sigma,\sigma'\in\mathbb{C}^2}} \frac{2\pi e^2}{\Omega(\mu^2+q^2)}\mathbf{c}_{\mathbf{k}+\mathbf{q}\,\sigma}^\dagger \mathbf{c}_{\mathbf{k}'-\mathbf{q}\,\sigma'}^\dagger \mathbf{c}_{\mathbf{k}'\,\sigma'}\mathbf{c}_{\mathbf{k}\,\sigma}.
\end{aligned}
$$

The $(\mathbf{q}=0)$-term is special,

(190)
$$
\begin{aligned}
\sum_{\substack{\mathbf{k},\mathbf{k}\in\Lambda \\ \sigma,\sigma'\in\mathbb{C}^2}} \frac{2\pi e^2}{\Omega\mu^2}\mathbf{c}_{\mathbf{k}\,\sigma}^\dagger \mathbf{c}_{\mathbf{k}'\,\sigma'}^\dagger \mathbf{c}_{\mathbf{k}'\,\sigma'}\mathbf{c}_{\mathbf{k}\,\sigma} &= \sum_{\substack{\mathbf{k},\mathbf{k}'\in\Lambda \\ \sigma,\sigma'\in\mathbb{C}^2}} \frac{2\pi e^2}{\Omega\mu^2}\mathbf{c}_{\mathbf{k}\,\sigma}^\dagger \mathbf{c}_{\mathbf{k}\,\sigma}\left(\mathbf{c}_{\mathbf{k}'\,\sigma'}^\dagger \mathbf{c}_{\mathbf{k}\,\sigma} - \delta_{\mathbf{k}\mathbf{k}'}\delta_{\sigma\sigma'}\right) \\
&= \frac{2\pi e^2}{\mu^2\Omega}(N^2 - N).
\end{aligned}
$$

The $\frac{N^2 e^2}{\mu}$-terms all add up to zero while the second term $-\frac{2\pi N e^2}{\mu^2\Omega} \xrightarrow[\Omega\to\infty]{} 0$, thus reducing the system's Hamiltonian to

(191)
$$
\boxed{\mathbf{H} = \sum_{\substack{\mathbf{k}\in\Lambda \\ \mathbf{q}\neq 0 \\ \sigma\in\mathbb{C}^2}} \frac{\hbar^2 k^2}{2m_e}\mathbf{c}_{\mathbf{k}\sigma}^\dagger \mathbf{c}_{\mathbf{k}\sigma} + \sum_{\substack{\mathbf{k},\mathbf{k}',\mathbf{q}\in\Lambda \\ \sigma,\sigma'\in\mathbb{C}^2}} \frac{2\pi e^2}{\Omega q^2}\mathbf{c}_{\mathbf{k}+\mathbf{q}\,\sigma}^\dagger \mathbf{c}_{\mathbf{k}'-\mathbf{q}\,\sigma'}^\dagger \mathbf{c}_{\mathbf{k}'\,\sigma'}\mathbf{c}_{\mathbf{k}\,\sigma},}
$$

where $\mu = 0$ since the $(q = 0)$-singularity has been removed. The first term is the Hamiltonian for the Sommerfeld gas and where the second term is a perturbation. Notice tha the ground state has been expressed as

a product of single particle wavefunctions and not as an antisymmetrized linear combination of such terms, since the antisymmetrization has already been taken care of by the creation and annihilation operator commutators. Then, the jellium ground-state is given by

$$|\Omega\rangle = \prod_{\mathbf{k}=0}^{k=k_F} |\mathbf{k}\uparrow\rangle \langle \mathbf{k}\downarrow|.$$

The Wigner crystal. In the small-density limit, the Coulomb energy become dominant and therefore it is more appropriate to start from the ground state of the interaction term than from that of the kinetic energy. Wigner argued that the free electron-like Fermi sphere ground state could become unstable to an insulating lattice of localized electrons when the electron density becomes sufficiently small. This leads to a purely classical problem. Namely, to calculate the lowest-energy configuration of charged particles immersed into a homogeneous background of opposite charge. This problem was first addressed by Wigner in 1934 as that of an inverted alkali metal, and he argued that at low enough densities the electrons would form a crystal. A consistent theory has of course to take into account the kinetic energy, leading to zero-point fluctuations of the electrons around their equilibrium positions. As the lattice constant decreases, these fluctuations become more and more important, until the Wigner crystal melts. Numerical simulations indicate this happens for $r_s \sim 100$. In fact, a two-dimensional Wigner crystal with a triangular structure has actually been observed for a very low-density electron system ($r_s \sim 10^4$) dispersed over the surface of liquid helium.

The first-order energy correction is given by

$$E^{(1)} = \langle\Omega| \mathbf{H}' |\Omega\rangle = \langle\Omega| \sum_{\substack{\mathbf{k},\mathbf{k}',\mathbf{q}\in\Lambda \\ \sigma,\sigma'\in\mathbb{C}^2}} \frac{2\pi e^2}{\Omega q^2} c^\dagger_{\mathbf{k}+\mathbf{q}\,\sigma} c^\dagger_{\mathbf{k}'-\mathbf{q}\,\sigma'} c_{\mathbf{k}'\,\sigma'} c_{\mathbf{k}\,\sigma} |\Omega\rangle.$$

For this matrix-elements to be non-zero, the states $\mathbf{k}+\mathbf{q}$, $\mathbf{k}'-\mathbf{q}$, $\mathbf{k}'$ and $\mathbf{k}$ must be contained in the ground-state wavefunction product. This means that the only terms that do not vanish will be those that do not change $|\Omega\rangle$'s occupation numbers, therefore

(192)
$$\begin{cases} \mathbf{k} = \mathbf{k}+\mathbf{q}, \quad \mathbf{k}' = \mathbf{k}'-\mathbf{q} \quad \mathbf{q}=0 \\ \mathbf{k}' = \mathbf{k}+\mathbf{q}, \quad \mathbf{k} = \mathbf{k}'-\mathbf{q} \end{cases}$$

This can be interpreted via figure 2. The first configuration implies no scattering since $\mathbf{q}=0$. The remaining term, the exchange scattering term, is represented by the second diagram in figure 2 and depicts the case where the particle in $|\mathbf{k}\rangle$-state is scattered into the $|\mathbf{k}'\rangle$-state and vice-versa. The correction too E_0 is the exchange energy, given by

(193)
$$\begin{aligned} E^{(1)} &= \langle\Omega| \sum_{\substack{\mathbf{k},\mathbf{k}',\mathbf{q}\in\Lambda \\ \sigma\in\mathbb{C}^2}} \frac{2\pi e^2}{\Omega q^2} c^\dagger_{\mathbf{k}+\mathbf{q}\,\sigma} c^\dagger_{\mathbf{k}'-\mathbf{q}\,\sigma} c_{\mathbf{k}'\,\sigma} c_{\mathbf{k}\,\sigma} |\Omega\rangle \\ &= \langle\Omega| \sum_{\substack{\mathbf{k},\mathbf{q}\in\Lambda \\ \sigma\in\mathbb{C}^2}} \frac{2\pi e^2}{\Omega q^2} c^\dagger_{\mathbf{k}+\mathbf{q}\,\sigma} c^\dagger_{\mathbf{k}\sigma} c_{\mathbf{k}+\mathbf{q}\,\sigma} c_{\mathbf{k}\sigma} |\Omega\rangle \\ &= \langle\Omega| \sum_{\substack{\mathbf{k},\mathbf{q}\in\Lambda \\ \sigma\in\mathbb{C}^2}} \frac{2\pi e^2}{\Omega q^2} (-N_{\mathbf{k}+\mathbf{q}\,\sigma} - N_{\mathbf{k}\sigma}) |\Omega\rangle \Rightarrow \varepsilon^{(1)}(\mathbf{k}) = -\frac{2\pi e^2}{\Omega} \sum_{\mathbf{q}} \frac{N_{\mathbf{k}+\mathbf{q}}}{q^2} \end{aligned}$$
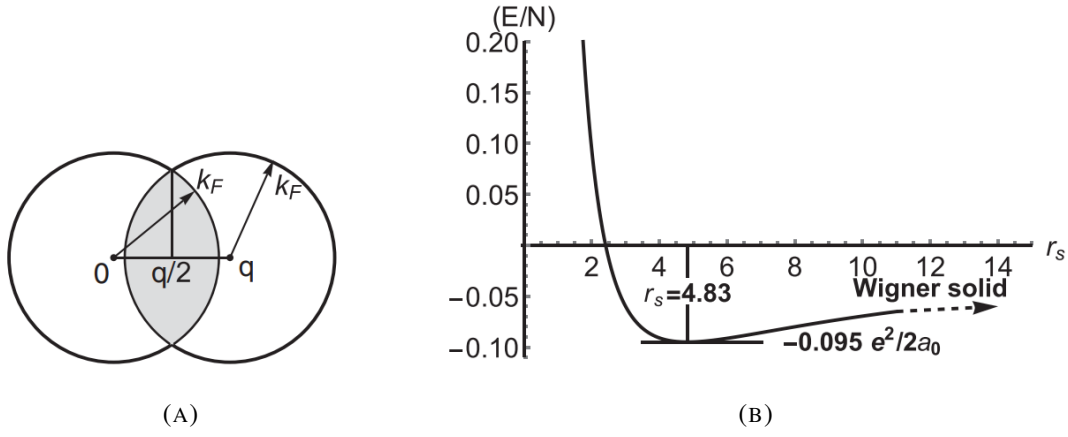
FIGURE 1. Fig. A: Geometry for the **k**-integration is shown for arbitrary, but fixed, **q**, which is just the intersection of two spheres.
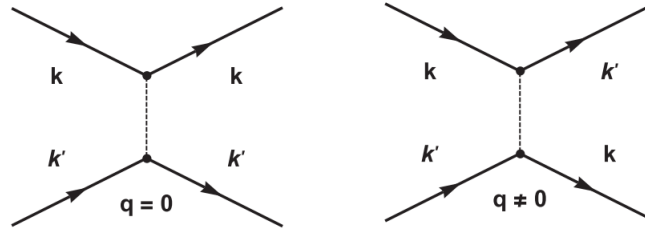Fig. B: Ground-state energy per electron, including first-order correction, for the jellium model.

FIGURE 2. Diagrams for the jellium sea. The dashed line represents the interaction.

(194)
$$\Rightarrow E^{(1)} = -\frac{4\pi e^2}{\Omega} \frac{\Omega}{(2\pi)^6} \int_{\mathbb{R}^6} d\mathbf{k} d\mathbf{q} \, \frac{1}{q^2} \Theta(k_F - |\mathbf{k} + \mathbf{q}|)\Theta(k_F - k) \text{ with the integration region shown in figure 1 (left)}$$
$$= -\frac{4\pi e^2 \Omega}{(2\pi)^6} \int_{\mathbb{R}^3} \frac{d\mathbf{q}}{q^2} \int_{\mathbb{R}^3} d\mathbf{p} \, \Theta\left(k_F - \left|\mathbf{p} + \frac{1}{2}\mathbf{q}\right|\right)\Theta\left(k_F - \frac{1}{2}\mathbf{q}\right)$$
$$= -\frac{e^2}{2a_0} N \frac{3}{2\pi} \left(\frac{9\pi}{4}\right)^{\frac{1}{3}} \frac{1}{r_s}.$$

This results, shown in **??** (right), shows that the electron gas is stable when the repulsive Coulomb interaction is turned on. No external confinement potential is therefore needed to hold the electron gas in the ion jellium together. There exists an optimal density n^*, or interparticle distance r_s^*, which minimizes the energy and furthermore yields an $E^* < 0$. The negative exchange energy overcome the positive kinetic energy. This method of treating the electron gas is a special case of the Hartree-Fock approximation and is the simplest way to take into account interactions. The interesting point to note is that in this approximation the energy is reduced below that of the Sommerfeld gas, which is a manifestation of the exchange hole.

Hartree-Fock approximation as a Mean-Field Theory.

Consider the electronic Hamiltonian given by equation (189). The goal is to write an effective two-body Hamiltonian, which takes into account the average effects of the interactions. Therefore, the four-Fermi interaction is replaced with a sum of all possible two-body terms

$$\langle \mathbf{c}_1^\dagger \mathbf{c}_2^\dagger \mathbf{c}_3 \mathbf{c}_4 \rangle_{\text{MFT}} = -\langle \mathbf{c}_1^\dagger \mathbf{c}_3 \rangle \mathbf{c}_2^\dagger \mathbf{c}_4 - \langle \mathbf{c}_2^\dagger \mathbf{c}_4 \rangle \mathbf{c}_1^\dagger \mathbf{c}_3 + \langle \mathbf{c}_1^\dagger \mathbf{c}_4 \rangle \mathbf{c}_2^\dagger \mathbf{c}_3 + \langle \mathbf{c}_2^\dagger \mathbf{c}_3 \rangle \mathbf{c}_1^\dagger \mathbf{c}_4,$$

(195)

$$\text{where } \langle \mathbf{c}_{\mathbf{k}\sigma}^\dagger \mathbf{c}_{\mathbf{k}'\sigma'} \rangle = \mathbf{n}_{\mathbf{k}\sigma} \delta_{\mathbf{k}\mathbf{k}'}^{\sigma\sigma'}$$

The mean field terms can be understood as the average number of particles $\mathbf{n}_{\mathbf{k}\sigma}$ in the $|\mathbf{k}\sigma\rangle$-state, which will be weighted with the two-body interaction $V(\mathbf{q})$ to give the average interaction due to all other interaction terms. Then, the four-Fermi interaction in the mean-field approximation reads

(196)

$$\mathcal{V}_{\text{HF}} = \frac{1}{2} \sum_{\substack{\mathbf{k}\mathbf{k}',\mathbf{q}\in\Lambda \\ \sigma\sigma'\in\mathbb{C}^2}} V(\mathbf{q}) \Big[- \langle \mathbf{c}_{\mathbf{k}\sigma}^\dagger \mathbf{c}_{\mathbf{k}'\sigma'} \rangle \mathbf{c}_{\mathbf{k}'-\mathbf{q}\,\sigma'}^\dagger \mathbf{c}_{\mathbf{k}-\mathbf{q}\sigma} - \langle \mathbf{c}_{\mathbf{k}'-\mathbf{q}\,\sigma'}^\dagger \mathbf{c}_{\mathbf{k}-\mathbf{q}\sigma} \rangle \mathbf{c}_{\mathbf{k}\sigma}^\dagger \mathbf{c}_{\mathbf{k}'\sigma'}$$

$$+ \langle \mathbf{c}_{\mathbf{k}\sigma}^\dagger \mathbf{c}_{\mathbf{k}-\mathbf{q}\sigma} \rangle \mathbf{c}_{\mathbf{k}'-\mathbf{q}\,\sigma'}^\dagger \mathbf{c}_{\mathbf{k}'\sigma'} + \langle \mathbf{c}_{\mathbf{k}'-\mathbf{q}\,\sigma'}^\dagger \mathbf{c}_{\mathbf{k}'\sigma'} \rangle \mathbf{c}_{\mathbf{k}\sigma}^\dagger \mathbf{c}_{\mathbf{k}-\mathbf{q}\sigma} \Big]$$

$$= - \sum_{\substack{\mathbf{k}\mathbf{q}\in\Lambda \\ \sigma\in\mathbb{C}^2}} V(\mathbf{q}) \langle \mathbf{c}_{\mathbf{k}\sigma}^\dagger \mathbf{c}_{\mathbf{k}\sigma} \rangle \mathbf{c}_{\mathbf{k}-\mathbf{q}\,\sigma}^\dagger \mathbf{c}_{\mathbf{k}-\mathbf{q}\,\sigma} + V(0) \sum_{\substack{\mathbf{k}\mathbf{k}'\in\Lambda \\ \sigma\sigma'\in\mathbb{C}^2}} \langle \mathbf{c}_{\mathbf{k}\sigma}^\dagger \mathbf{c}_{\mathbf{k}\sigma} \rangle \mathbf{c}_{\mathbf{k}'\sigma'}^\dagger \mathbf{c}_{\mathbf{k}'\sigma'}$$

$$= \sum_{\substack{\mathbf{k}\in\Lambda \\ \sigma\in\mathbb{C}^2}} \Big(-\sum_{\mathbf{q}\in\Lambda} \mathbf{n}_{\mathbf{k}+\mathbf{q}\sigma} V(\mathbf{q}) + nV(0) \Big) \mathbf{c}_{\mathbf{k}\sigma}^\dagger \mathbf{c}_{\mathbf{k}\sigma},$$

$$\Rightarrow \mathbf{H}^{\text{HF}} = \sum_{\substack{\mathbf{k}\in\Lambda \\ \sigma\in\mathbb{C}^2}} \varepsilon_{\mathbf{k}}^{\text{HF}} \mathbf{c}_{\mathbf{k}\sigma}^\dagger \mathbf{c}_{\mathbf{k}\sigma}, \text{ where} \quad \begin{aligned} \varepsilon_{\mathbf{k}}^{\text{HF}} &= \varepsilon_{\mathbf{k}} + \sum_{\substack{\mathbf{k}'\in\Lambda \\ \sigma'\in\mathbb{C}^2}} [V(0) - \delta_{\sigma\sigma'} V(\mathbf{k}-\mathbf{k}')] \mathbf{n}_{\mathbf{k}'\sigma'} \\ &= \varepsilon_{\mathbf{k}} + V(0)N - \sum_{\substack{\mathbf{k}'\in\Lambda \\ \sigma'\in\mathbb{C}^2}} V(\mathbf{k}-\mathbf{k}') \mathbf{n}_{\mathbf{k}'\sigma'}. \end{aligned}$$

Note that the Hartree or direct Coulomb term, which represents the average interaction energy of the electron $\mathbf{k}\sigma$ with all the other electrons in the system, is a constant; when summed over all electron it cancels exactly with the constant arising from the sum of the self-energy of the positive background and the interaction energy of the electron gas with that of the background. The Fock, or exchange term, is a momentum-dependent shift.

Problems with the Hartree-Fock Theory

Although we argued that the Hartree-Fock approximation becomes a better approximation in the high-density limit for electron interacting via the Coulomb attraction, it never becomes exact. The Hartree-Fock energy correction is given by

$$\varepsilon^{(1)}(\mathbf{k}) = \frac{e^2 k_F}{4\pi^2} \Big[\frac{k_F^2 - k^2}{k k_F} \log \Big| \frac{k_F + k}{k_F - k} \Big| + 2 \Big]$$

The first term in parentheses has a logarithmic divergence in slope $k = kf$. This means that while the energy shift might be small compared to the Fermi energy, the Fermi velocity $v_F = \nabla_{\mathbf{k}} \varepsilon(\mathbf{k})$, contains a non-finite term. This problem can be traced back to the long-range nature of the Coulomb force. Two electrons at large distances $\mathbf{x} - \mathbf{x}'$ do not feel the full $\frac{1}{|\mathbf{x}-\mathbf{x}'|}$-interaction, but a screened version due to the presence of the intervening medium, namely the electron gas, rearranges itself to cancel out the long-range part of V.

Electron Interaction in Second-order Perturbation Theory

Trying to improve upon the first-order result by going to second-order perturbation theory yields a disastrous results for the matrix elements diverge. According to second-order perturbation theory,

(197)

$$\frac{E^{(2)}}{N} = \frac{1}{N} \sum_{\substack{|\psi\rangle\in\mathbb{H} \\ |\psi\rangle\neq|\Omega\rangle}} \frac{\langle\Omega| \mathbf{V}_c |\psi\rangle \langle\psi| \mathbf{V}_c |\Omega\rangle}{E^{(0)} - E_\psi}.$$
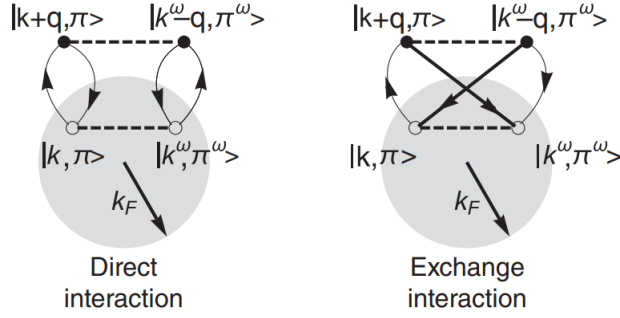
FIGURE 3. The only two possible processes in second-order perturbation theory for two states, $|\mathbf{k}_1, \sigma_1\rangle$ and $|\mathbf{k}_2, \sigma_2\rangle$, in the Fermi sea. The direct process gives a divergent contribution to $\frac{E}{N}$ while the exchange process gives a finite contribution.

This, combined with the momentum-conserving Coulomb interaction, yields intermediate states where two particles are injected out of the Fermi sphere. From such an intermediate state, $|\Omega\rangle$, is restored by putting the excited electrons back into the holes they left behind. Therefore, the are only two types of possible processes: direct interaction and the exchange process, shown in figure 3.

1) In the <u>Direct interaction</u>, it turns out that the direct interaction gives a divergent contribution $E_{\text{dir}}^{(2)}$ to $E^{(2)}$ due to the singular behaviour of the Coulomb interaction at small $\mathbf{q}$-momentum transfers. The $(|\psi\rangle = |\Omega\rangle)$-constraint leads to the second p

$$|\psi\rangle = \Theta(|\mathbf{k}_1 + \mathbf{q}| - k_F)\Theta(|\mathbf{k}_2 - \mathbf{q}| - k_F)\Theta(k_F - |\mathbf{k}_2|)\Theta(k_F - |\mathbf{k}_1|)c_{\mathbf{k}_1+\mathbf{q}}^\dagger c_{\mathbf{k}_2-\mathbf{q}}^\dagger c_{\mathbf{k}_2} c_{\mathbf{k}_1} |\Omega\rangle.$$

To restore $|\Omega\rangle$, the same $\mathbf{q}$-momentum transfer must be involved in both $\langle\Omega| \mathbf{V_c} |\text{-}\rangle$ and $\langle\psi| \mathbf{V_c} |\Omega\rangle$, and writing $\mathbf{V}(\mathbf{q}) = \frac{4\pi e^2}{\mathbf{q}^2}$ yields

$$E_{\text{dir}}^{(2)} = \frac{1}{4\Omega^2} \sum_{\mathbf{q}\in\Lambda} \sum_{\substack{\mathbf{k}\mathbf{k}'\in\Lambda \\ \sigma\sigma'\in\mathbb{C}^2}} \frac{\mathbf{V}(\mathbf{q})^2}{E^{(0)} - E_\nu} \Theta(|\mathbf{k}_1 + \mathbf{q}| - k_F)\Theta(|\mathbf{k}_2 - \mathbf{q}| - k_F)\Theta(k_F - |\mathbf{k}_2|)\Theta(k_F - |\mathbf{k}_1|)$$

(198)

Given that $\mathbf{V}(\mathbf{q})^2 \approx q^{-4} \Rightarrow E^{(0)} - E_\psi \overset{q\to 0}{\approx} \mathbf{k}_1^2 + \mathbf{k}_2^2 - (\mathbf{k}_1 + \mathbf{q})^2 - (\mathbf{k}_2 - \mathbf{q})^2 \approx q$

$$\sum_{\mathbf{q}\in\Lambda} \sum_{\substack{\mathbf{k}\mathbf{k}'\in\Lambda \\ \sigma\sigma'\in\mathbb{C}^2}} \cdots \Theta(|\mathbf{k}_2 - \mathbf{q}| - k_F)\Theta(k_F - |\mathbf{k}_2|)\Theta(k_F - |\mathbf{k}_1|) \overset{q\to 0}{\approx} q$$

$$\Rightarrow E_{\text{dir}}^{(2)} = \int_0 dq q^2 \frac{1}{q^4} \frac{1}{q} qq = \int_0 \frac{dq}{q} = \log q|_0 \to \infty.$$

2) The exchange process does not lead to a divergence since, in this case, the momentum transfer in the excitation part is $\mathbf{q}$, but in the relaxation part it is $\mathbf{k}_2 - \mathbf{k}_1 - \mathbf{q}$. Thus $\mathbf{V}^2(\mathbf{q})$ is replaced by $\mathbf{V}^2(\mathbf{q})\mathbf{V}(\mathbf{k}_2 - \mathbf{k}_1 - \mathbf{q}) \propto q^{-2}$ for $\mathbf{q} \to 0$, which is less singular than $\mathbf{V}^2(\mathbf{q}) \propto q^{-4}$.

Physically, the energy of an electron gas must be finite. The only hope for rescue lies in regularization of the divergent behaviour by taking higher-order perturbation terms into account. Consider now the ground-state energy of the jellium model of a three-dimensional electron gas. In the $(r_s < 1)$-regime, this energy can be expressed in terms of a power-series in r_s, as

$$E_0 = \frac{K}{r_s^2}[1 + br_s + cr_s^2 + \cdots],$$

where K, a, b, c are constants. This is not exactly right, since first-order perturbation theory gives a term of the (br_s)-form in the series. But, considering second-order perturbation theory, one finds a contribution that diverges like $\int_0 \frac{dq}{q}$, where q is the transfer-momentum in the Fourier transform $\mathbf{V}_q$ of the Coulomb interaction, $\mathbf{V}_q \approx q^{-2}$. Thus, there is in fact a logarithmic divergence from the lower limit 0 of the momentum transfers. This divergence is associated with the long range of the Coulomb interaction. Furthermore, if one examines higher-order terms in the perturbation series, one finds that they diverge even more strongly.

On physical grounds, however, the energy of the interacting electron gas must be finite and well-defined, with no phase transitions occurring as the repulsive interactions are turned on. This failure of standard perturbation theory appears to signal that the energy does not have a standard power-series expansion in r_s. In 1957, Gell-Mann and Bruckner resolved this issue by developing many-body perturbation theory[16], in which the summation of the divergent terms in the series an infinite number of terms) is performed before doing the momentum integrals, thus arriving to a finite and well-defined results. This is called the resummation of the perturbation series. They found that there was a term in the series for E_0 that is $\propto \log r_s$, and is non-analytic at $r_s = 0$. This holds so long the interaction doesn't cause any drastic changes to the system, such as a phase transition.

3.5. **Random Phase Approximation.** In the second quantization formalism, the fermionic field operators may be written in plane-wave basis as

$$\Psi(\mathbf{x}, t) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{k} \in \Lambda} e^{i\mathbf{k} \cdot \mathbf{x}} \mathbf{c_k}, \quad \Psi^\dagger(\mathbf{x}, t) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{k} \in \Lambda} e^{-i\mathbf{k} \cdot \mathbf{x}} \mathbf{c_k^\dagger}.$$

The particle density $\rho(\mathbf{x})$ at point $\mathbf{x}$ is given by

$$\rho(\mathbf{x}) = \sum_{i \in \mathbb{N}} \delta(\mathbf{x} - \mathbf{x}_i) \Rightarrow \rho(\mathbf{x}) = \int_{\mathbb{R}^3} d\mathbf{x} \, \Psi^\dagger(\mathbf{x}) \Big(\sum_{i \in \mathbb{N}} \delta(\mathbf{x} - \mathbf{x}_i) \Big) \Psi(\mathbf{x}) = \sum_{i \in \mathbb{N}} \Psi^\dagger(\mathbf{x}) \Psi(\mathbf{x}).$$

Its Fourier transform is given by $\rho_{\mathbf{q}} = \sum_{\mathbf{k} \in \Lambda} \mathbf{c_{k+q}^\dagger} \mathbf{c_k}$. Its physical meaning is straightforward: it gives rise to particle-hole pairs charge fluctuations around the Fermi surface. Since $\rho^\dagger = \rho$, it follows that $\rho_{\mathbf{q}}^\dagger = \rho_{-\mathbf{q}}$.[17]

Now, the following question naturally arises. Under what conditions an interacting system, with Hamiltonian

$$\mathbf{H} = \sum_{\mathbf{k} \in \Lambda} \varepsilon_{\mathbf{k}} \mathbf{c_k^\dagger} \mathbf{c_k} + \frac{1}{2} \sum_{\mathbf{q} \in \Lambda} \mathbf{V_q} (\rho_{\mathbf{q}} \rho_{-\mathbf{q}} - N),$$

and ground-state $|\Omega\rangle$, $\mathbf{H} |\Omega\rangle = \varepsilon_0 |\Omega\rangle$, can support density excitations associated with $\rho_{\mathbf{q}}$? Namely, these excitations are

$$\mathbf{H} \rho_{\mathbf{q}}^\dagger |\Omega\rangle = (\varepsilon_0 + \hbar \omega_{\mathbf{q}}) \rho_{\mathbf{q}}^\dagger |\Omega\rangle$$

The previous equation may be rewritten as

---

[16]Many-body perturbation expansions without diagrams. I. Normal states. Behnam Farid §

[17]For example, the following four-body term may be recast as

$$\sum_{\mathbf{k}, \mathbf{k}' \in \Lambda} \mathbf{c_{k-q}^\dagger} \mathbf{c_{k'+q}^\dagger} \mathbf{c_{k'}} \mathbf{c_k} = - \sum_{\mathbf{k}' \in \Lambda} \mathbf{n_{k'}} + \rho_{\mathbf{q}} \rho_{-\mathbf{q}'}.$$

$$[\mathbf{H}, \rho_\mathbf{q}^\dagger] = \hbar\omega_\mathbf{q}\rho_\mathbf{q}^\dagger \Rightarrow \left[\sum_{\mathbf{k}\in\Lambda} \varepsilon_\mathbf{k} \mathbf{c}_\mathbf{k}^\dagger \mathbf{c}_\mathbf{k} + \frac{1}{2}\sum_{\mathbf{q}\in\Lambda} \mathbf{V}_\mathbf{q}(\rho_\mathbf{q}\rho_{-\mathbf{q}} - N), \rho_\mathbf{q}^\dagger \right] = \hbar\omega_\mathbf{q}\rho_\mathbf{q}^\dagger$$

$$\left[\sum_{\mathbf{k}\in\Lambda} \varepsilon_\mathbf{k} \mathbf{c}_\mathbf{k}^\dagger \mathbf{c}_\mathbf{k}, \rho_\mathbf{q}^\dagger \right] + \frac{1}{2}\left[\sum_{\mathbf{q}\in\Lambda} \mathbf{V}_\mathbf{q}(\rho_\mathbf{q}\rho_{-\mathbf{q}} - N), \rho_\mathbf{q}^\dagger \right] =$$

$$\sum_{\mathbf{k}\in\Lambda} \varepsilon_\mathbf{k} \left[\mathbf{c}_\mathbf{k}^\dagger \mathbf{c}_\mathbf{k}, \rho_\mathbf{q}^\dagger \right] + \frac{1}{2}\sum_{\mathbf{q}\in\Lambda} \mathbf{V}_\mathbf{q} \left[\rho_\mathbf{q}\rho_{-\mathbf{q}}, \rho_\mathbf{q}^\dagger \right] =$$

(199)

$$\sum_{\mathbf{k},\mathbf{\ell}\in\Lambda} \varepsilon_\mathbf{k} \left[\mathbf{c}_\mathbf{k}^\dagger \mathbf{c}_\mathbf{k}, \mathbf{c}_\mathbf{q} \mathbf{c}_{\mathbf{\ell}+\mathbf{q}}^\dagger \right] =$$

$$\sum_{\mathbf{k},\mathbf{\ell}\in\Lambda} \varepsilon_\mathbf{k} \left(\mathbf{c}_\mathbf{k}^\dagger [\mathbf{c}_\mathbf{k}, \mathbf{c}_\mathbf{q}] \mathbf{c}_{\mathbf{\ell}+\mathbf{q}}^\dagger + [\mathbf{c}_\mathbf{k}^\dagger, \mathbf{c}_\mathbf{q}] \mathbf{c}_\mathbf{k} \mathbf{c}_\mathbf{k}^\dagger + \mathbf{c}_\mathbf{q} \mathbf{c}_\mathbf{k}^\dagger [\mathbf{c}_\mathbf{k}, \mathbf{c}_{\mathbf{\ell}+\mathbf{q}}^\dagger] + \mathbf{c}_\mathbf{q} [\mathbf{c}_\mathbf{k}^\dagger, \mathbf{c}_{\mathbf{\ell}+\mathbf{q}}^\dagger] \mathbf{c}_\mathbf{k} \right) =$$

$$\sum_{\mathbf{k}\in\Lambda} (\varepsilon_{\mathbf{k}+\mathbf{q}} - \varepsilon_\mathbf{k}) \mathbf{c}_{\mathbf{k}+\mathbf{q}}^\dagger \mathbf{c}_\mathbf{k} =$$

If $\rho_\mathbf{q}^\dagger$ actually creates an excitation of energy $\hbar\omega$, then

(200)

$$[\mathbf{H}, [\mathbf{H}, \rho_\mathbf{q}^\dagger]] = (\hbar\omega)^2 \rho_\mathbf{q}^\dagger$$

$$\Rightarrow (\hbar\omega)^2 \rho_\mathbf{q}^\dagger = \sum_{\mathbf{k}\in\Lambda} (\varepsilon_{\mathbf{k}+\mathbf{q}} - \varepsilon_\mathbf{k})[\mathbf{H}, \mathbf{c}_{\mathbf{k}+\mathbf{q}}^\dagger \mathbf{c}_\mathbf{k}] = \sum_{\mathbf{k}\in\Lambda} (\varepsilon_{\mathbf{k}+\mathbf{q}} - \varepsilon_\mathbf{k})^2 \mathbf{c}_{\mathbf{k}+\mathbf{q}}^\dagger \mathbf{c}_\mathbf{k}$$

$$+ \sum_{\mathbf{k},\mathbf{q}'\in\Lambda} (\varepsilon_{\mathbf{k}+\mathbf{q}} - \varepsilon_\mathbf{k}) \frac{V_{\mathbf{q}'}}{2} \left[(\mathbf{c}_{\mathbf{k}+\mathbf{q}-\mathbf{q}'}^\dagger \mathbf{c}_\mathbf{k} - \mathbf{c}_{\mathbf{k}+\mathbf{q}}^\dagger \mathbf{c}_{\mathbf{k}+\mathbf{q}'})\rho_{\mathbf{q}'}^\dagger + \rho_{\mathbf{q}'}(\mathbf{c}_{\mathbf{k}+\mathbf{q}+\mathbf{q}'}^\dagger \mathbf{c}_\mathbf{k} - \mathbf{c}_{\mathbf{k}+\mathbf{q}}^\dagger \mathbf{c}_{\mathbf{k}-\mathbf{q}'}) \right]$$

$$= \sum_{\mathbf{k}\in\Lambda} \left[\frac{\hbar^2}{2m}(2\mathbf{k}\cdot\mathbf{q} + \mathbf{q}^2) \right]^2 \mathbf{c}_{\mathbf{k}+\mathbf{q}}^\dagger \mathbf{c}_\mathbf{k} + \sum_{\mathbf{k},\mathbf{q}'\in\Lambda} \frac{V_{\mathbf{q}'}}{2} \left[\frac{\hbar^2}{2m}(2\mathbf{k}\cdot\mathbf{q} + \mathbf{q}^2)(\mathbf{c}_{\mathbf{k}+\mathbf{q}-\mathbf{q}'}^\dagger \mathbf{c}_\mathbf{k} - \mathbf{c}_{\mathbf{k}+\mathbf{q}}^\dagger \mathbf{c}_{\mathbf{k}+\mathbf{q}'}) \right.$$

$$\left. - \frac{\hbar^2}{2m}(2\mathbf{k}\cdot\mathbf{q} + \mathbf{q}^2)(\mathbf{c}_{\mathbf{k}+\mathbf{q}+\mathbf{q}'}^\dagger \mathbf{c}_\mathbf{k} - \mathbf{c}_{\mathbf{k}+\mathbf{q}}^\dagger \mathbf{c}_{\mathbf{k}-\mathbf{q}'}) \right]$$

$$= \sum_{\mathbf{k}\in\Lambda} \left[\frac{\hbar^2}{2m}(2\mathbf{k}\cdot\mathbf{q} + \mathbf{q}^2) \right]^2 \mathbf{c}_{\mathbf{k}+\mathbf{q}}^\dagger \mathbf{c}_\mathbf{k} + \sum_{\mathbf{q}'\in\Lambda} \frac{V_{\mathbf{q}'}}{2} \left[\sum_{\mathbf{k}\in\Lambda} \frac{\hbar^2}{2m}(2\mathbf{k}\cdot\mathbf{q} + \mathbf{q}^2)(\mathbf{c}_{\mathbf{k}+\mathbf{q}-\mathbf{q}'}^\dagger \mathbf{c}_\mathbf{k} - \mathbf{c}_{\mathbf{k}+\mathbf{q}}^\dagger \mathbf{c}_{\mathbf{k}+\mathbf{q}'}) \right.$$

$$\left. - \sum_{\mathbf{k}\in\Lambda} \frac{\hbar^2}{2m}(2\mathbf{k}\cdot\mathbf{q} + \mathbf{q}^2)(\mathbf{c}_{\mathbf{k}+\mathbf{q}+\mathbf{q}'}^\dagger \mathbf{c}_\mathbf{k} - \mathbf{c}_{\mathbf{k}+\mathbf{q}}^\dagger \mathbf{c}_{\mathbf{k}-\mathbf{q}'}) \right]$$

$$= \sum_{\mathbf{k}\in\Lambda} \left[\frac{\hbar^2}{2m}(2\mathbf{k}\cdot\mathbf{q} + \mathbf{q}^2) \right]^2 \mathbf{c}_{\mathbf{k}+\mathbf{q}}^\dagger \mathbf{c}_\mathbf{k}$$

$$+ \sum_{\mathbf{q}'\in\Lambda} \frac{V_{\mathbf{q}'}}{2} \left[\sum_{\mathbf{k}\in\Lambda} \frac{\hbar^2}{2m}(2\mathbf{k}\cdot\mathbf{q} + \mathbf{q}^2)\mathbf{c}_{\mathbf{k}+\mathbf{q}-\mathbf{q}'}^\dagger \mathbf{c}_\mathbf{k} - \frac{\hbar^2}{2m}(2\mathbf{k}\cdot\mathbf{q} + \mathbf{q}^2)\mathbf{c}_{\mathbf{k}+\mathbf{q}}^\dagger \mathbf{c}_{\mathbf{k}+\mathbf{q}'} \right.$$

$$\left. - \sum_{\mathbf{k}\in\Lambda} \frac{\hbar^2}{2m}(2\mathbf{k}\cdot\mathbf{q} + \mathbf{q}^2)\mathbf{c}_{\mathbf{k}+\mathbf{q}+\mathbf{q}'}^\dagger \mathbf{c}_\mathbf{k} - \frac{\hbar^2}{2m}(2\mathbf{k}\cdot\mathbf{q} + \mathbf{q}^2)\mathbf{c}_{\mathbf{k}+\mathbf{q}}^\dagger \mathbf{c}_{\mathbf{k}-\mathbf{q}'} \right]$$

$$= \sum_{\mathbf{k}\in\Lambda} \left[\frac{\hbar^2}{2m}(2\mathbf{k}\cdot\mathbf{q} + \mathbf{q}^2) \right]^2 \mathbf{c}_{\mathbf{k}+\mathbf{q}}^\dagger \mathbf{c}_\mathbf{k} \qquad \text{let } \mathbf{k}' = \mathbf{k} + \mathbf{q}', \ \mathbf{k}'' = \mathbf{k} - \mathbf{q}'$$

(201)

$$+ \sum_{\mathbf{q}'\in\Lambda} \frac{V_{\mathbf{q}'}}{2} \left[\sum_{\mathbf{k}\in\Lambda} \frac{\hbar^2}{2m}(2\mathbf{k}\cdot\mathbf{q} + \mathbf{q}^2)\mathbf{c}_{\mathbf{k}+\mathbf{q}-\mathbf{q}'}^\dagger \mathbf{c}_\mathbf{k} - \sum_{\mathbf{k}'\in\Lambda} \frac{\hbar^2}{2m}(2(\mathbf{k}'-\mathbf{q}')\cdot\mathbf{q} + \mathbf{q}^2)\mathbf{c}_{\mathbf{k}'+\mathbf{q}-\mathbf{q}'}^\dagger \mathbf{c}_{\mathbf{k}'} \right.$$

$$\left. - \sum_{\mathbf{k}\in\Lambda} \frac{\hbar^2}{2m}(2\mathbf{k}\cdot\mathbf{q} + \mathbf{q}^2)\mathbf{c}_{\mathbf{k}+\mathbf{q}+\mathbf{q}'}^\dagger \mathbf{c}_\mathbf{k} - \sum_{\mathbf{k}''\in\Lambda} \frac{\hbar^2}{2m}(2(\mathbf{k}''+\mathbf{q}')\cdot\mathbf{q} + \mathbf{q}^2)\mathbf{c}_{\mathbf{k}''+\mathbf{q}+\mathbf{q}'}^\dagger \mathbf{c}_{\mathbf{k}''} \right]$$
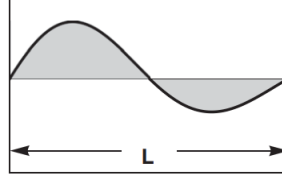
FIGURE 4. Consider a length-L box. Then, ρ_q measures the quantity approximately equal to the difference in the particle number in the two halves of the container.

$$(202) \quad (\hbar\omega)^2\rho_{\mathbf{q}}^\dagger = \sum_{\mathbf{k}\in\Lambda}\left[\frac{\hbar^2}{2m}(2\mathbf{k}\cdot\mathbf{q}+\mathbf{q}^2)\right]^2 c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}} + \sum_{\mathbf{q}'\in\Lambda}\frac{V_{\mathbf{q}'}}{2}\frac{\hbar^2\mathbf{q}\cdot\mathbf{q}'}{m}\left(\rho_{\mathbf{q}'-\mathbf{q}}\rho_{\mathbf{q}'}^\dagger - \rho_{-\mathbf{q}'}^\dagger\rho_{-\mathbf{q}-\mathbf{q}'}\right).$$

Now, in order to study equation (202), note that ρ_0 is actually much larger than all the other components. It is just the average particle density in the system. Consider a box of length-L, shown in figure 4, containing a large number of electrons of ρ_0 average density, then ρ_0's first Fourier component is

$$\int_{\Omega\subset\mathbb{R}^3} d\mathbf{x}\,\rho(\mathbf{x})e^{\frac{2i\pi x}{L}},$$

will be approximately equal to the difference between the particle number in the left and right hand sides of the box, the gray area. The number difference will then be small, compared to the total number and so ρ_0 will prevail in the summation over $\mathbf{q}'$. Since the ($\mathbf{q}'=0$)-term is omitted, ρ_0 will only appear when $\mathbf{q}'=\pm\mathbf{q}$. The neglect of all other terms, when $\mathbf{q}'\neq\pm\mathbf{q}$ is known as the **random phase approximation** [18].

Plasma oscillations
Consider the case where $\mathbf{q}$ is small enough that, for the moment, the first term in equation (202) may be neglected and the second term approximated via RPA. In that case,

$$(\hbar\omega)^2\rho_{\mathbf{q}}^\dagger = \frac{4\pi e^2}{q^2}\frac{\hbar^2 q^2}{m}(\rho_0\rho_{\mathbf{q}}^\dagger + \rho_{\mathbf{q}}^\dagger\rho_0) \rightarrow \omega^2 = \frac{4\pi e^2\rho_0}{m},$$

which gives just the classical plasma frequency ω_P. The contribution of the first term may be approximated by setting $k\sim k_F$ and $\langle 2\mathbf{k}_F\cdot\mathbf{q}\rangle >> \mathbf{q}^2$. Averaging over all angles yields

$$(\mathbf{k}_F\cdot\mathbf{q})^2\rangle \simeq \frac{1}{4\pi}4\pi^2 k_F^2 q^2\int_{-1}^{1}dx\,x^2 = 2\pi k_F^2 q^2,$$

which yield the plasmon dispersion relation

$$\omega^2(\mathbf{q}) = \omega_P^2 + \frac{2\pi\hbar^2 k_F^2}{3m^2}q^2 \simeq \omega_P + \frac{\pi v_F^2}{3\omega_P}q^2.$$

Thus, the relevant long-wavelength excitations are collective motions of the electron gas. Bohm and Pines suggested that electrons act as interacting particles through a bare Coulomb potential $\frac{4\pi e^2}{q^2}$ only for length scales less than an effective screening length, which may be taken to be the Thomas-Fermi distance, $\sim q_{\text{TF}}^{-1}$. Below q_{TF}, the interaction contribute only to the plasma oscillations.

---

[18] An alternative interpretation for the RPA is as follows. Consider the first-quantized version of the density operator $\rho_{\mathbf{q}\in\Lambda} = \frac{1}{\Omega}\sum_{n\in\mathbb{N}}e^{i\mathbf{q}\cdot\mathbf{x}_n}$, where $\mathbf{x}_n$ is the n-th's particles position. Then, $\rho_{\mathbf{q}\pm\mathbf{q}'}$ may be written in first quantized form as

$$\rho_{\mathbf{q}\pm\mathbf{q}'}\frac{1}{\Omega}\sum_{n\in\mathbb{N}}e^{i(\mathbf{q}\pm\mathbf{q}')\cdot\mathbf{x}_n}.$$

Then, if the electrons are statistically randomly distributed in the system, the preceding sum only gives a finite result if $\mathbf{q}=\mp\mathbf{q}'$, since the sum involves random phases that add destructively. In RPA, equation (202) is approximated by keeping only terms with $\mathbf{q}=\mp\mathbf{q}'$.

3.6. **Green Functions for Many-Body systems.** Green function techniques play a fundamental role in the treatment of many-body systems, for many properties of a many body system -eg. ground state energy, excitation spectra, response functions, electric and magnetic susceptibilities and thermodynamic quantities- can be derived from many-body Green functions.

Consider the ground state $|\Omega\rangle$ of a system. At some time t', a perturbation is enacted upon the system, said perturbation represented by a local operator $\mathbf{O}^\dagger(\mathbf{x}, t)$, creating an initial state $\mathbf{O}^\dagger(\mathbf{x}', t') |\Omega\rangle$. Thus, the system has been deviated from the initial some by some local excitations, a disturbance, and then this disturbance will evolve in space and time. Such a scenario describes typical local measurements done on a quantum many-particle system. Since an explicit time dependence is necessary for this formalism, the Heisenberg picture is useful. To determine physically interesting quantities, the quantum-mechanical amplitude to find the system in state $\mathbf{O}^\dagger(\mathbf{x}', t') |\Omega\rangle$ at time $t > t'$ namely,

$$\mathcal{A} = \langle\Omega| \mathbf{O}(\mathbf{x}, t)\mathbf{O}^\dagger(\mathbf{x}', t') |\Omega\rangle .$$

Many quantum-mechanical operators may be written in bilinear forms of the field operators themselves, $\hat{\Psi}(\mathbf{x}, t)$. However, it turns out that correlation functions of the field operators are related to experiments such as photoemission and tunnelling spectroscopies, which involve extraction or injection of a single electron. These field correlation functions can be identified as propagators which satisfy the Green function equation.

3.6.1. *The One-Particle Green Function of Many-Body Systems.*
Non-interacting Particle Propagator. Consider a particle in free space described by a single-particle time-independent Hamiltonian $\mathbf{H}_1$. Its eigenstates and eigenenergies are

$$\mathbf{H}_1 |\phi_n\rangle = \varepsilon_n |\phi_n\rangle .$$

In general, if a particle is in the $|\phi\rangle_i$-state, it will remain in that state forever. Instead, consider a system in a generic state $|\psi_{\text{trial}}\rangle$. If the trial state is created at time t', the wave-function at a later time t is given by

$$|\psi(t)\rangle = e^{-\frac{i\mathbf{H}(t-t')}{\hbar}} |\psi_{\text{trial}}\rangle = \sum_n |\psi_n\rangle e^{-\frac{i\varepsilon_n(t-t')}{\hbar}} \bar{\phi}_n |\psi_{\text{trial}}\rangle$$

4. LINEAR RESPONSE THEORY

Linear response theory is an extremely widely used concept in physics, stating that the response to a weak external perturbation is proportional to the perturbation, and therefore the quantity of interest is the proportionality constant. The physical question to ask is thus: supposing some perturbation H', what is the measured consequence for an observable quantity $\mathbf{A}$. In other words, what is $\langle \mathbf{A} \rangle$ to linear order in H'?

Among the numerous physical application of the linear response formalism, one can mention charge and spin susceptibilities of eg. electron systems due to external electric or magnetic fields. Responses to external mechanical forces or vibrations can also be calculated using the same formalism.

The general Kubo formula. Consider a quantum system described by a time independent Hamiltonian $\mathbf{H}_0$ in thermodynamic equilibrium. This means that an expectation value of a physical quantity, described by the operator $\mathbf{A}$, which can be evaluated as

(203)
$$\langle \mathbf{A} \rangle = \frac{1}{\mathcal{Z}_0} \mathrm{Tr}\ _\mathbb{H} \rho_0 \mathbf{A} = \frac{1}{\mathcal{Z}_0} \sum_{n \in \Lambda} \langle n | \mathbf{A} | n \rangle e^{-\beta E_n} \quad \text{where } \{|n\rangle\}_{n \in \Lambda} \text{ is a complete}$$
$$\rho_0 = e^{-\beta \mathbf{H}_0} = \sum_{n \in \Lambda} |n\rangle \langle n| e^{-\beta E_n}, \qquad \text{set of eigenstates.}$$

Suppose now that at some time, $t = t_0$, an external perturbation is applied to the system, driving it out of equilibrium. The perturbation is described by an additional time dependent term in the Hamiltonian

(204)
$$\mathbf{H}(t) = \mathbf{H}_0 + \mathbf{H}'(t)\theta(t - t_0)$$

Now, the interest lies in finding the expectation value of the $\mathbf{A}$ operator at times t greater that t_0-. In order to do so, the time evolution of the density matrix must be found, or equivalently the time evolution of the eigenstates of the unperturbed Hamiltonian. Once $|n(t)\rangle$ is found, the time-dependent expectation value can be found as

(205)
$$\langle \mathbf{A}(t) \rangle = \frac{1}{\mathcal{Z}_0} \mathrm{Tr}\ _\mathbb{H} \rho(t) \mathbf{A} = \frac{1}{\mathcal{Z}_0} \sum_{n \in \Lambda} \langle n(t) | \mathbf{A} | n(t) \rangle e^{-\beta E_n}$$

(206)
$$\rho_0 = e^{-\beta \mathbf{H}_0} = \sum_{n \in \Lambda} |n(t)\rangle \langle n(t)| e^{-\beta E_n},$$

The physical idea behind this expression is as follows. The initial states of the system are distributed according to the usual Boltzmann distribution $\frac{e^{-\beta E_{0n}}}{\mathcal{Z}_t}$. At later times, the system is described by the same distribution of states but the states are now time-dependent and they have evolved according to the new Hamiltonian. The time dependence of the states $|n(t)\rangle$ is governed by the Schrödinger equation. Since $\mathbf{H}'$ is regarded to be a small perturbation, the interaction picture representation is suitable for this setting. In this representation, the time dependence is given by

(207)
$$|n(t)\rangle = e^{-i\mathbf{H}_0 t} |n(t)\rangle_I = e^{-i\mathbf{H}_0 t} \mathcal{U}(t, t_0) |\hat{n}(t_0)\rangle,$$

where by definition $|\hat{n}(t_0)\rangle = e^{i\mathbf{H}_0 t_0} |n(t_0)\rangle = |n\rangle$.

Up to linear order in $\mathbf{H}'$, the time evolution operator $\mathcal{U}(t, t_0)$ can be written

$$\mathcal{U}(t, t_0) = \mathbb{1} - i \int_{t_0}^{t} dt' \mathbf{H}'(t') + \mathcal{O}(\mathbf{H}'^2)$$

Then, using this in the time-dependent expectation value yields

(208)

$$\langle \mathbf{A}(t) \rangle = \frac{1}{\mathcal{Z}_0} \sum_{n \in \Lambda} \langle n(t) | \mathbf{A} | n(t) \rangle e^{-\beta E_n}$$

$$= \frac{1}{\mathcal{Z}_0} \sum_{n \in \Lambda} \langle n(t_0) | \left(\mathbb{1} - i \int_{t_0}^{t} dt' \mathbf{H}'(t') \right) e^{-i\mathbf{H}_0 t_0} \mathbf{A} e^{i\mathbf{H}_0 t_0} \left(\mathbb{1} - i \int_{t_0}^{t} dt' \mathbf{H}'(t') \right) | n(t_0) \rangle e^{-\beta E_n} + \mathcal{O}(\mathbf{H}'^2)$$

$$= \frac{1}{\mathcal{Z}_0} \sum_{n \in \Lambda} \langle n(t_0) | \mathbf{A} | n(t_0) \rangle - \frac{i}{\mathcal{Z}_0} \int_{t_0}^{t} dt' \sum_{n \in \Lambda} e^{-\beta E_n} \left(e^{-i\mathbf{H}_0 t_0} \mathbf{A} e^{i\mathbf{H}_0 t_0} \mathbf{H}'(t') - \mathbf{H}'(t') e^{-i\mathbf{H}_0 t_0} \mathbf{A} e^{i\mathbf{H}_0 t_0} | n(t_0) \rangle \right)$$

$$= \langle \mathbf{A} \rangle_0 + \int_{t_0}^{t} \frac{dt'}{i} \langle [\mathbf{A}(t), \mathbf{H}'(t')] \rangle_0,$$

where the brackets $\langle \rangle_0$ means an equilibrium average with respect to the Hamiltonian $\mathbf{H}$. This is in fact a remarkable and very useful result since the inherently non-equilibrium quantity $\langle \mathbf{A}(t) \rangle$ has been expressed as a correlation function of the system in equilibrium. The physical reason for this is that the interaction between excitations created in the non-equilibrium state is an effect to second order in the weak perturbation, hence not included in the linear response.

The correlation function is the retarded correlation function, which can be rewritten as the difference between $\langle \mathbf{A}(t) \rangle$ and $\langle \mathbf{A} \rangle_0$ ie.

(209)

$$\delta \langle \mathbf{A}(t) \rangle = \int_{t_0}^{\infty} dt' \mathcal{C}_{AH'}^{R}(t, t') e^{-\eta(t-t')} \text{ where } \mathcal{C}_{AH'}^{R}(t, t') = -i\theta(t - t') \langle [\mathbf{A}(t), \mathbf{H}'(t')] \rangle_0,$$

which is the Kubo formula. This expresses the linear response to a perturbation $\mathbf{H}'$. Note that the factor $e^{-\eta(t-t')}$, with an infinitesimal positive parameter η, has been included to force the response at time t due to the influence of $\mathbf{H}'$ at time t' to decay when $t >> t'$. At the end of the calculation, the limit $\eta \to 0^+$,. This is so since the retarded effect of a perturbation must decrease in time[19].

Kubo fromula in the frequency domain. It is often convenient to express the response to an external disturbance in the frequency domain via Fourier transformations[20]. Therefore, consider the perturbation Hamiltonian $\mathbf{H}'$, which can be rewritten in terms of its Fourier components

(211)
$$\mathbf{H}'(t) = \int_{\Omega \subset \mathbb{R}} \frac{d\omega}{2\pi} e^{-i\omega t} \mathbf{H}'_\omega,$$

such that the retarded correlation function becomes

(212)
$$\mathcal{C}_{AH'}^{R}(t, t') = \int_{\mathbb{R}} \frac{d\omega}{2\pi} e^{-i\omega t'} \mathcal{C}_{AH'_\omega}^{R}(t - t') \quad \begin{array}{l} \text{since } \langle [\mathbf{A}(t), \mathbf{H}'(t')_{\omega'}] \rangle_0 \text{ only depends} \\ \text{on the difference between } t \text{ and } t'. \end{array}$$

Therefore, inserting this result into the Kubo formula yields

---

[19]The other, the advanced correlation function is non-physical and must be ditched. The other one, the retarded correlation function, decreases exponentially with time, the exponential factor thus picks out the physically relevant solution by introducing an artificial relaxation mechanism.

[20]Consider the L^1-space, ie. the space of all integrable functions on the real line. Then the Fourier transform and the Fourier-anti transform can be defined as

(210)
$$\mathcal{F}[f(t)](\omega) = f(\omega) = \int_{\mathbb{R}} dt e^{i\omega t} f(t) \text{ and } \mathcal{F}^{-1}[f(\omega)](t) = f(t) = \int_{\mathbb{R}} \frac{d\omega}{2\pi} e^{-i\omega t} f(\omega)$$

$$\delta\langle \mathbf{A}(t)\rangle = \int_{t_0}^{\infty} dt' \mathcal{C}_{AH'_\omega}^{R}(t,t') e^{-\eta(t-t')}$$

$$= \int_{\mathbb{R}} dt' \int_{\mathbb{R}} \frac{d\omega}{2\pi} e^{-i\omega t} e^{-i(\omega+i\eta)(t'-t)} \mathcal{C}_{AH'_\omega}^{R}(t-t')$$

(213)
$$= \int_{\mathbb{R}} \frac{d\omega}{2\pi} e^{-?^i\omega t}\left(\int_{\mathbb{R}} d(t'-t) e^{-i(\omega+i\eta)(t'-t)} \mathcal{C}_{AH'_\omega}^{R}(t-t')\right)$$

$$= \int_{\mathbb{R}} \frac{d\omega}{2\pi} e^{-i\omega t} \mathcal{C}_{AH'_\omega}^{R}(\omega;\eta)$$

which can be inverted to yield the final result in the frequency domain

$$\int_{\mathbb{R}} dt e^{i\nu t}\delta\langle \mathbf{A}(t)\rangle = \int_{\mathbb{R}} dt e^{i\nu t}\int_{\mathbb{R}} \frac{d\omega}{2\pi} e^{-i\omega t}\mathcal{C}_{AH'_\omega}^{R}(\omega;\eta)$$

(214)
$$\langle \mathbf{A}_\nu\rangle = \int_{\mathbb{R}} \frac{d\omega}{2\pi}\int_{\mathbb{R}} dt e^{i(\nu-\omega)t}\mathcal{C}_{AH'_\omega}^{R}(\omega;\eta)$$

$$= \int_{\mathbb{R}} d\omega\, \delta(\nu-\omega)\mathcal{C}_{AH'_\omega}^{R}(\omega;\eta)$$

$$= \mathcal{C}_{AH'_\nu}^{R}(\nu;\eta)$$

(215)
$$\boxed{\Rightarrow \delta\langle \mathbf{A}_\omega\rangle = \mathcal{C}_{AH'_\omega}^{R}(\omega) \text{ with } \mathcal{C}_{AH'_\omega}^{R}(\omega) = \int_{\mathbb{R}} dt\, e^{i\omega t} e^{-\eta t}\mathcal{C}_{AH'_\omega}^{R}(t)},$$

where the infinitesimal η parameter is incorporated in order to ensure the correct physical result, ie. the retarded response function decays at $t >> 1$.

Kubo formula for conductivity. Consider a system of charged particles, eg. electrons, which is subjected t an external electromagnetic field. The electromagnetic field induces a current and the conductivity is the linear response coefficient. In the general case, the conductivity is a non-local quantity in both time and space, such that the electric current $\mathbf{J}_e$ at some point $\mathbf{x}$ at time t depends on the electric field at points $\mathbf{y}$ at times t', eg.[21]

(216)

$$J_e^\alpha(x^\mu) = \int_{\Omega\subset\mathbb{R}^4} dx^\mu \sum_{\alpha\beta} \sigma_\beta(x^\mu,y^\mu)E^\beta(y^\mu),$$

where $\sigma_{\alpha\beta}(x^\mu,y^\mu)$ is the conductivity tensor, which describes the current response in the $\mathbf{e}_\alpha$-direction to an applied electric field in the $\mathbf{e}_\beta$-direction.

The electric field $\mathbf{E}$ is given by the electric potential ϕ_{ext} and the vector potential $\mathbf{A}_{\text{ext}}$ as $\mathbf{E}(x^\mu) = -\partial_\mu A^\mu$. For electrons, the current density can be written as $\mathbf{J}_e = -e\langle \mathbf{J}\rangle$. The perturbing term in the Hamiltonian due to the external electromagnetic field is given by the coupling of the electrons to both the scalar potential and the vector potential. Then, upto linear order in the external potential,

$$\mathbf{H}_{\text{ext}} = -e\int_{\mathbb{R}^3} d\mathbf{x}\, J_\mu(\mathbf{x})A_{\text{ext}}^\mu(x^\mu).$$

Let, $\mathbf{A}_0$ denote the vector potential in the equilibrium ie. prior to the onset of the perturbation $\mathbf{A}_0(x^\mu$ and let $\mathbf{A}_0(x^\mu)$ denote the total vector potential. Then,

$$\mathbf{A}(x^\mu) = \mathbf{A}_0(x^\mu) + \mathbf{A}_{\text{ext}}(x^\mu).$$

The current operator can be decomposed in two components, the diamagnetic and the paramagnetic terms, as follows

---

[21]Note that this section's mathematical treatment is not Lorentz-covariant. This is, the spacetime is treated as $\mathbb{R}^4$ with the metric being $g_{\mu\nu} = \delta_{\mu\nu}$

(217)
$$\mathbf{J} = \mathbf{J}^{\nabla}(\mathbf{x}) + \frac{e}{m}\mathbf{A}(\mathbf{x})\rho(\mathbf{x}).$$

For simplicity and using gauge invariance, the external electric potential can be set to zero. The conductivity is most easily expressed in the frequency domain via a Fourier transformation of the perturbation. Since

(218)
$$\partial_t \xrightarrow{F} -i\omega \quad \text{then } \mathbf{A}_{\text{ext}}(\mathbf{x},\omega) \xrightarrow{F} \frac{1}{i\omega}\mathbf{E}_{\text{ext}}(\mathbf{x},\omega) \ \Rightarrow \mathbf{H}_{\text{ext},\omega} = \frac{e}{i\omega}\int_{\Omega\subset\mathbb{R}^3} d\mathbf{x}\, \mathbf{J}(\mathbf{x})\cdot\mathbf{E}_{\text{ext}}(\mathbf{x},\omega).$$

In order to exploit the frequency domain formulation of linear response theory, it is desirable to find the corresponding formula for the conductivity tensor in frequency-space. The conductivity tensor is a property of the equilibrium system and can thus only depend on time differences $\sigma_{\alpha\beta}(x^\mu, y^\mu) = \sigma_{\alpha\beta}(\mathbf{x},\mathbf{y},t-t')$. The frequency transform of the conductivity yields

(219)
$$J_e^\alpha(\mathbf{x},\omega) = \int_{\Omega\subset\mathbb{R}^3} d\mathbf{y}\, \sum_\beta \sigma_{\alpha\beta}(\mathbf{x},\mathbf{y},\omega)E^\beta(\mathbf{y},\omega).$$

Now, given that he external perturbation, written in frequency space, is already linear in the external potential $\mathbf{E}_{\text{ext}}$, and given that the interest lies only on terms proportional to said perturbation, the conductivity can be rewritten as

5. MICROSCOPIC THEORY OF CONVENTIONAL SUPERCONDUCTIVITY

From **Advanced Quantum Condensed Matter Physics, One-Body, Many-Body and Topological perspectives. Michael El-Batanouny**

.