

Theory & Notes

1. SPIN CHAIN MODELS

1.1. 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-dimensional 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 is said to be unsatisfied and so the spins are frustrated, yielding completely new properties different from the unfrustrated model.

1.1.1. The partition function and correlators. Consider a general spin chain as a dynamical system. Its spins can take any value in a given configuration space, which is a finite-dimensional manifold \mathcal{M} . Therefore, assuming that \mathcal{M} is a smooth manifold, the tangent space at $|s\rangle$ is denoted by $T_{|s\rangle}\mathcal{M}$. Then, the cotangent space at $|s\rangle$ is defined as the dual space of $T_{|s\rangle}\mathcal{M}$,

$$(1) \quad T_{|s\rangle}^*\mathcal{M} = (T_{|s\rangle}\mathcal{M})^*.$$

Concretely, elements of the cotangent space are linear functionals on $T_{|s\rangle}\mathcal{M}$ that is, every element $\langle s| \in T_{|s\rangle}^*\mathcal{M}$ is a linear map $\langle s| : T_{|s\rangle}\mathcal{M} \rightarrow \mathbb{C}$.

Then, for for each value $|s\rangle \in \mathcal{M}$, the "momentum" $\langle s|$ of the system would take values in the cotangent space $T_{|s\rangle}^*\mathcal{M}$ of that space. Thus, the phase space is naturally represented here by the cotangent bundle

$$T^*\mathcal{M} = \{(|s\rangle, \langle s|) : |s\rangle \in \mathcal{M} \text{ and } \langle s| \in T_{|s\rangle}^*\mathcal{M}\}$$

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 $n \in \mathcal{M}$, in terms of said configuration's energy and system's temperature. In other words,

$$(2) \quad \mathbb{P} : T^* \mathcal{M} \rightarrow \mathbb{R}_{[0,1]} \text{ where } \mathbb{P}(n) = \frac{e^{-\beta E_n}}{\mathcal{Z}},$$

where \mathcal{Z} is the partition function, defined by the requirement that the probabilities sum to one, that is

$$\mathcal{Z} = \sum_{n \in T^* \mathcal{M}} e^{-\beta E_n}.$$

Note that if the degrees of freedom take on continuous values, or the model is defined in the continuum, then this sum is replaced by an integral. 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.

1.2. Quantum Spin Systems.

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 $\phi : \mathbb{V} \rightarrow \mathbb{V}$ be a homomorphism 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

$$(3) \quad \mathbf{H}_{\text{XXX}} = -J \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j = -J \sum_{i=1}^N \left(\frac{1}{2} (\sigma_i^+ \sigma_{i+1}^- + \sigma_i^- \sigma_{i+1}^+) + \frac{1}{4} \sigma_i^z \sigma_{i+1}^z \right),$$

where in the ferromagnetic case the diagonal terms in the Hamiltonian favour aligned spins, while in the antiferromagnetic case, the diagonal terms favour antialignment. 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

$$(4) \quad \mathbf{H}_{\text{XXZ}} = - \sum_{\langle jk \rangle} (J_{\perp} (\mathbf{S}_j^x \mathbf{S}_k^x + \mathbf{S}_j^y \mathbf{S}_k^y) + J_z \mathbf{S}_j^z \mathbf{S}_k^z) = - \sum_{j=1}^N \left(J_{\perp} (\sigma_j^+ \sigma_{j+1}^- + \sigma_j^- \sigma_{j+1}^+) + \frac{\Delta}{2} \sigma_j^z \sigma_{j+1}^z \right),$$

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 $\Delta \rightarrow \pm\infty$ -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,

$\Delta \rightarrow -\infty$ 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 $\mathbf{S}_j^a \rightarrow -\mathbf{S}_j^a$.

- 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. $\uparrow \rightarrow \downarrow$ 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,

$$(5) \quad \begin{aligned} \text{triplet} : & |\uparrow\uparrow\rangle, \frac{1}{\sqrt{2}}(|\downarrow\uparrow\rangle + |\uparrow\downarrow\rangle), |\downarrow\downarrow\rangle, \text{ with } \lambda_i = -\frac{J}{4} \\ \text{singlet} : & \frac{1}{\sqrt{2}}(|\downarrow\uparrow\rangle - |\uparrow\downarrow\rangle), \text{ with } \mu_i = +\frac{J}{4} \end{aligned}$$

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 periodic 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

$$(6) \quad \begin{aligned} & \mathcal{T} : \mathfrak{A}_i \rightarrow \mathfrak{A}_j \\ & \mathcal{T}^{-1} \vec{S}_i \mathcal{T} = \vec{S}_{i+1} \end{aligned} \quad \text{where } \mathfrak{A}_i, \mathfrak{A}_j \simeq \mathfrak{su}(2) \forall i, j.$$

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 translation 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

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

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

$$(7) \quad \begin{aligned} |A\rangle &= \frac{1}{\sqrt{2}} \left(|\uparrow\downarrow\uparrow\downarrow\rangle + e^{-i\pi k} |\downarrow\uparrow\downarrow\uparrow\rangle \right), & \text{with the hamiltonian} \\ |B\rangle &= \frac{1}{2} \left(|\uparrow\uparrow\downarrow\downarrow\rangle + |\downarrow\downarrow\uparrow\uparrow\rangle + \cdots \right) & \text{on these} \\ & & k = 0 \text{ and } k = \pi \\ & & \text{states being} \end{aligned} \quad -J \begin{pmatrix} -1 & \sqrt{2} \cos\left(\frac{\pi k}{2}\right) \\ \sqrt{2} \cos\left(\frac{\pi k}{2}\right) & 0 \end{pmatrix}.$$

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

$$(8) \quad \begin{aligned} & \text{quintuplet : } 1 \\ & \text{triplets : } \cos(\pi k) (k \neq 0) \\ & \text{singlets : } -2, 0 \text{ for } k = 0, \pi \text{ respectively.} \end{aligned}$$

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.

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

$$(9) \quad \mathbf{H} = J \sum_{i=1}^L (\mathbf{S}_i^x \mathbf{S}_{i+1}^x + \mathbf{S}_i^y \mathbf{S}_{i+1}^y) - \lambda \sum_{j=1}^L \mathbf{S}_j^z,$$

which describes interacting spins in a one-dimensional chain, with periodic boundary conditions. (9)'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 (9) 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.

$$\begin{aligned} \mathbf{S}_j^\pm &= \mathbf{S}_j^x \pm i\mathbf{S}_j^y \Rightarrow \begin{aligned} \mathbf{S}_j^x &= \frac{1}{2}(\mathbf{S}_j^+ + \mathbf{S}_j^-), \\ \mathbf{S}_j^y &= \frac{1}{2i}(\mathbf{S}_j^+ - \mathbf{S}_j^-). \end{aligned} \end{aligned}$$

Then, the XX-Hamiltonian can be re-written as

$$(10) \quad \mathbf{H} = \frac{J}{2} \sum_{i=1}^L (\mathbf{S}_i^+ \mathbf{S}_{i+1}^- + \mathbf{S}_i^- \mathbf{S}_{i+1}^+) - \lambda \sum_{j=1}^L \mathbf{S}_j^z$$

where the interacting terms in the first summation, flip neighboring spins if said spins are anti-aligned¹. In addition, the XX-Hamiltonian has a total magnetization symmetry, since the Hamiltonian given by (10) commutes with the magnetization operator.

Now, we use a Jordan-Wigner transformation whereby the spin operators are mapped to fermionic operators, as follows

$$\begin{aligned} \mathbf{S}_j^z &= f_j^\dagger f_j - \frac{1}{2} \\ \mathbf{S}_j^- &= \exp\left(i\pi \sum_{\ell=1}^{j-1} f_\ell^\dagger f_\ell\right) f_j \\ \mathbf{S}_j^+ &= \exp\left(-i\pi \sum_{\ell=1}^{j-1} f_\ell^\dagger f_\ell\right) f_j^\dagger \end{aligned} \quad (12)$$

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),

¹In 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.

$$(11) \quad (\mathbf{S}_1^+ \mathbf{S}_2^- + \mathbf{S}_1^- \mathbf{S}_2^+) |\downarrow\downarrow\rangle = |\downarrow\downarrow\rangle \text{ and } (\mathbf{S}_1^+ \mathbf{S}_2^- + \mathbf{S}_1^- \mathbf{S}_2^+) |\uparrow\uparrow\rangle = |\uparrow\uparrow\rangle$$

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

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

which is not problematic, since it accounts for all L -lattice sites. Let

$$(13) \quad \mathbf{S}_L^+ \mathbf{S}_1^- = \mathcal{Q} f_L^\dagger f_1 \text{ and } \mathbf{S}_L^- \mathbf{S}_1^+ = \mathcal{Q} f_1^\dagger f_L,$$

then (10) can be rewritten as

$$(14) \quad \mathbf{H} = \frac{J}{2} \sum_{i=1}^{L-1} \left(f_j^\dagger f_{j+1} + f_{j+1}^\dagger f_j \right) - \lambda \sum_{j=1}^L \left(f_j^\dagger f_j - \frac{1}{2} \right) + \frac{J}{2} \mathcal{Q} (f_L^\dagger f_1 + f_1^\dagger f_L),$$

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 (14) can also have periodic boundary conditions, thus yielding

$$(15) \quad \mathbf{H} = \frac{J}{2} \sum_{i=1}^L \left(f_j^\dagger f_{j+1} + f_{j+1}^\dagger f_j \right) - \lambda \sum_{j=1}^L \left(f_j^\dagger f_j - \frac{1}{2} \right) + \frac{J}{2} (\mathcal{Q} - 1) (f_L^\dagger f_1 + f_1^\dagger f_L),$$

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

$$(16) \quad \mathbf{H} = \frac{J}{2} \sum_{i=1}^L \left(f_j^\dagger f_{j+1} + f_{j+1}^\dagger f_j - \lambda f_j^\dagger f_j \right) + \frac{\lambda L}{2},$$

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

$$(17) \quad f_j = \frac{1}{\sqrt{L}} \sum_{\substack{k=2\pi m/L \\ m \in \mathbb{Z}_{[1,L]}}} e^{ijk} d_k, \quad f_j^\dagger = \frac{1}{\sqrt{L}} \sum_{\substack{k=2\pi m/L \\ m \in \mathbb{Z}_{[1,L]}}} e^{-ijk} d_k^\dagger,$$

with the inverse transformation given by

$$(18) \quad d_k = \frac{1}{\sqrt{L}} \sum_{j=1}^L e^{-ikj} f_j, \quad d_k^\dagger = \frac{1}{\sqrt{L}} \sum_{j=1}^L e^{ikj} f_j^\dagger.$$

Note that the d_k -operators follow the standard fermionic anticommutation 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, (16)'s terms are mapped as follows

$$\begin{aligned}
\sum_{j=1}^L f_j^\dagger f_{j+1} &= \sum_{j=1}^L \frac{1}{L} \sum_{k,q} e^{-ikj} e^{iq(j+1)} d_k^\dagger d_q = \sum_{j=1}^L \frac{1}{L} \sum_{k,q} e^{i(q-k)j} e^{iq} d_k^\dagger d_q \\
&= \sum_{k,q} \frac{1}{L} e^{iq} \delta_{qk} d_k^\dagger d_q = \sum_k e^{ik} d_k^\dagger d_k \\
\sum_{j=1}^L f_{j+1}^\dagger f_j &= \sum_{j=1}^L \frac{1}{L} \sum_{k,q} e^{-ik(j+1)} e^{iqj} d_k^\dagger d_q = \sum_{j=1}^L \frac{1}{L} \sum_{k,q} e^{i(q-k)j} e^{-ik} d_k^\dagger d_q \\
&= \sum_{k,q} \frac{1}{L} e^{-ik} \delta_{qk} d_k^\dagger d_q = \sum_k e^{-ik} d_k^\dagger d_k \\
\sum_{j=1}^L f_j^\dagger f_j &= \sum_{j=1}^L \frac{1}{L} \sum_{k,q} e^{-ikj} e^{iqj} d_k^\dagger d_q = \sum_{j=1}^L \frac{1}{L} \sum_{k,q} e^{i(q-k)j} d_k^\dagger d_q \\
&= \sum_{k,q} \frac{1}{L} \delta_{qk} d_k^\dagger d_q = \sum_k d_k^\dagger d_k
\end{aligned}$$

Therefore, using these identities, the new Hamiltonian is given by

$$\begin{aligned}
\mathbf{H} &= \frac{J}{2} \sum_k \left(e^{ik} d_k^\dagger d_k + e^{-ik} d_k^\dagger d_k - \lambda d_k^\dagger d_k \right) + \frac{\lambda L}{2} \\
(19) \quad &= \sum_k \left(J \cos k - \lambda \right) d_k^\dagger d_k + \frac{\lambda L}{2},
\end{aligned}$$

which can be rewritten as

$$(20) \quad \mathbf{H} = \sum_k \epsilon_k d_k^\dagger d_k + \frac{\lambda L}{2} \quad \begin{array}{l} \text{with the eigenvalues being given by } \epsilon_k = J \cos k - \lambda + \frac{\lambda L}{2} \\ \text{and the eigenvectors being given by } |E_n\rangle = \prod_n (d_k^\dagger)^n, \\ \text{with eigenvalue } E_n = \sum_n \epsilon_n \end{array}$$

where, $k = \frac{2\pi m}{L}$, $m = -\frac{L}{2} + 1, \dots, \frac{L}{2}$, 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.

$$(21) \quad \mathcal{N}_{jk} = \langle d_j^\dagger d_k \rangle_{\text{th}} = \frac{1}{1 + e^{\beta \epsilon_k + \mu}} \delta_{jk}$$

1.3.1. *Quantum Phase Transition.* The system's groundstate is defined such that

$$(22) \quad \langle E_G | d_k^\dagger d_k | E_G \rangle = 1 \text{ if } \epsilon_k \leq 0 \text{ and } \langle E_G | d_k^\dagger d_k | E_G \rangle = 0 \text{ if } \epsilon_k > 0.$$

Then the groundstate energy is given by

$$(23) \quad E_G = \sum_k \epsilon_k \langle E_G | d_k^\dagger d_k | E_G \rangle.$$

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

- when $\lambda < -J$, which yields a Fermi momentum $k_F = \pi$,
- when $|\lambda| \leq J$, which yields a Fermi momentum $k_F = \arccos\left(\frac{\lambda}{J}\right)$,
- when $\lambda > J$ which yields $k_F = 0$.

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

$$(24) \quad \begin{aligned} \langle E_G | d_k^\dagger d_k | E_G \rangle &= 1 & \langle E_G | d_k^\dagger d_k | E_G \rangle &= 0 \\ &\text{and} & & \\ \text{if } k \in [-\pi, -k_F] \cup [k_F, \pi], & & \text{if } k \in [-k_F, k_F]. \end{aligned}$$

The total magnetization operator is defined as the sum of the S_z operators, along the entire chain ie. $M^z = \sum_{j=1} S_j^z$. Under the Jordan-Wigner transformation, $S_j^z = f_j^\dagger f_j - \frac{1}{2}$. 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

$$(25) \quad M^z = \sum_k \left(d_k^\dagger d_k - \frac{1}{2} \right).$$

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

$$(26) \quad \langle E_G | M^z | E_G \rangle = \sum_{k=-\pi}^{-k_F} 1 + \sum_{k=k_F}^{\pi} 1 - \frac{1}{2} \sum_{k=-\pi}^{\pi} 1 = \frac{1}{2} \left(\sum_{k=-\pi}^{-k_F} 1 + \sum_{k=k_F}^{\pi} 1 - \sum_{k=-k_F}^{k_F} 1 \right),$$

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, $m^z = \frac{1}{L} M^z$. Now, as the chain gets progressively bigger, with more and more lattice sites, the summations in (26) can be approximated by Riemannian integrals, ie.

$$\sum_k f(k) = \frac{1}{\Delta k} \sum_k f(k) \Delta k \xrightarrow{L \rightarrow \infty} \frac{L}{2\pi} \int_{k \in \text{BZ}} dk f(k) \text{ where } \Delta k = \frac{2\pi}{L}.$$

Using this trick, (26) can be rewritten as

$$(27) \quad \langle m^z \rangle_G = \frac{1}{4\pi} \left(\int_{-\pi}^{-k_F} dk + \int_{k_F}^{\pi} dk - \int_{-k_F}^{k_F} dk \right) = \frac{1}{4\pi} \left(-k_F + \pi + \pi - k_F - k_F + k_F \right)$$

$$(28) \quad = \frac{1}{2} - \frac{k_F}{\pi},$$

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

$$(29) \quad \langle \mathbf{m}^z \rangle_G = \begin{cases} \frac{1}{2} - \frac{1}{\pi} \arccos\left(\frac{\lambda}{J}\right) & \text{if } |\lambda| \leq J \\ -\frac{1}{2} & \text{if } \lambda < J \\ \frac{1}{2} & \text{if } \lambda > J \end{cases}$$

Another interesting quantity is the total magnetic susceptibility, which can be written in terms of the total magnetization per lattice site as follows

$$(30) \quad \chi = \left(\frac{\partial \langle \mathbf{m}^z \rangle_G}{\partial \lambda} \right) \Rightarrow \chi = \begin{cases} \frac{1}{J\pi\sqrt{1-\frac{\lambda^2}{J^2}}} & \text{if } |\lambda| \leq J \\ 0 & \text{if } \lambda < J \\ 0 & \text{if } \lambda > J \end{cases},$$

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.

1.4. 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

$$(31) \quad \begin{aligned} \mathbf{H} &= J \sum_{n=1}^N \left(\vec{\mathbf{S}}_n \cdot \vec{\mathbf{S}}_{n+1} - \frac{1}{4} \mathbb{1}^{\otimes N} \right) \\ &= \frac{J}{4} \sum_{n=1}^N \left(\vec{\sigma}_n \cdot \vec{\sigma}_{n+1} - \frac{1}{4} \mathbb{1}^{\otimes N} \right). \end{aligned}$$

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 aligned 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 $\mathfrak{J} = \mathbb{1}_n / \mathbb{1}_k = \{k+1, \dots, n\}$. Let $A \in \text{GL}(n, \mathbb{C})$, which can be written in block notation as

$$(32) \quad A = \begin{pmatrix} A_{\mathcal{I}} & A_{\mathcal{I}\mathfrak{J}} \\ A_{\mathfrak{J}\mathcal{I}} & A_{\mathfrak{J}} \end{pmatrix} \begin{matrix} \mathbb{C}^k \\ \mathbb{C}^{n-k} \end{matrix} \text{ where } \begin{matrix} A_{\mathcal{I}} \in \text{GL}(k, \mathbb{C}), & A_{\mathcal{I}\mathfrak{J}} \in \text{GL}(k \rightarrow n-k, \mathbb{C}), \\ A_{\mathfrak{J}\mathcal{I}} \in \text{GL}(n-k \rightarrow k, \mathbb{C}), & A_{\mathfrak{J}} \in \text{GL}(n-k, \mathbb{C}), \end{matrix}$$

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 $\text{End}(\mathbb{C}^n)$ can be rewritten as a 2×2 -block-matrix, with its entries being matrices of a given dimension ie. ²

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

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

$$(33) \quad \mathbb{H} \simeq \bigotimes_{n=1}^N \mathbb{V}_n \text{ where } \mathbb{V}_n \simeq \mathbb{C}^2 \quad \begin{matrix} \text{To these individual Hilbert spaces, additional} \\ \text{auxiliary, non-physical, spaces } \mathbb{V}_a \simeq \mathbb{C}^2, \text{ can be added.} \end{matrix}$$

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

²For 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^α act on \mathbb{V}_n

$$(34) \quad \begin{aligned} \mathcal{L} : \mathbb{V}_n \otimes \mathbb{V}_a &\rightarrow \mathbb{V}_n \otimes \mathbb{V}_a \\ \mathcal{L}_{n,a} &= u(\mathbb{1}_n \otimes \mathbb{1}_a) + i \sum_{\alpha} \mathbf{S}_n^\alpha \otimes \sigma_a^\alpha, \quad \forall n \leq N \end{aligned}$$

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

complex-valued u -parameter is the spectral parameter.

Note that [equation \(34\)](#) can be rewritten as

$$(35) \quad \begin{aligned} \mathcal{L}_{n,a} &= u(\mathbb{1}_n \otimes \mathbb{1}_a) + i \sum_{\alpha} \mathbf{S}_n^\alpha \otimes \sigma_a^\alpha \\ &= \begin{pmatrix} u\mathbb{1}_n + i\mathbf{S}_n^z & i\mathbf{S}_n^x + \mathbf{S}_n^y \\ i\mathbf{S}_n^x - \mathbf{S}_n^y & u\mathbb{1}_n - i\mathbf{S}_n^z \end{pmatrix}_a = \begin{pmatrix} u\mathbb{1}_n + i\mathbf{S}_n^z & i\mathbf{S}_n^- \\ i\mathbf{S}_n^+ & u\mathbb{1}_n - i\mathbf{S}_n^z \end{pmatrix}_a, \end{aligned}$$

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

$$(36) \quad \mathcal{P}_{n,a} = \frac{1}{2} \left(\mathbb{1}_n \otimes \mathbb{1}_a + \sum_{\alpha} \sigma_n^\alpha \otimes \sigma_n^\alpha \right) = \frac{1}{2} \sigma^{n\mu} \sigma_\mu^\alpha$$

which ³ acts on both the physical and auxiliary systems.

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

$$(38) \quad \begin{aligned} \mathcal{L}_{n,a} &= u(\mathbb{1}_n \otimes \mathbb{1}_a) + i \sum_{\alpha} \mathbf{S}_n^\alpha \otimes \sigma_a^\alpha = u(\mathbb{1}_n \otimes \mathbb{1}_a) + \frac{i}{2} \sum_{\alpha} \sigma_n^\alpha \otimes \sigma_a^\alpha \\ &= u(\mathbb{1}_n \otimes \mathbb{1}_a) - \frac{i}{2} (\mathbb{1}_n \otimes \mathbb{1}_a) + \frac{i}{2} (\mathbb{1}_n \otimes \mathbb{1}_a) + \frac{i}{2} \sum_{\alpha} \sigma_n^\alpha \otimes \sigma_a^\alpha \\ &= \left(u - \frac{i}{2} \right) (\mathbb{1}_n \otimes \mathbb{1}_a) + \frac{i}{2} \left(\mathbb{1}_n \otimes \mathbb{1}_a + \sum_{\alpha} \sigma_n^\alpha \otimes \sigma_n^\alpha \right) \\ &= \left(u - \frac{i}{2} \right) \mathbb{1}_{n,a} + i \mathcal{P}_{n,a}. \end{aligned}$$

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

$$\begin{aligned} \mathcal{L}_{n,a_1}(u_1) : \mathbb{V}_n \otimes \mathbb{V}_{a_1} &\rightarrow \mathbb{V}_n \otimes \mathbb{V}_{a_1} \\ \mathcal{L}_{n,a_2}(u_2) : \mathbb{V}_n \otimes \mathbb{V}_{a_2} &\rightarrow \mathbb{V}_n \otimes \mathbb{V}_{a_2} \end{aligned} \quad \begin{array}{l} \text{both of which act on the physical quantum} \\ \text{state and on two different auxiliary spaces as well}^4. \end{array}$$

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

³Note that the permutation operator can be written as a 2×2 -matrix in the auxiliary space \mathbb{V}_a ,

$$(37) \quad \begin{aligned} \mathcal{P}_{n,a} &= \frac{1}{2} \left[\begin{pmatrix} \mathbb{1}_n & 0 \\ 0 & \mathbb{1}_a \end{pmatrix} + \sigma_n^x \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}_a + \sigma_n^y \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}_a + \sigma_n^z \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}_a \right] \\ &= \begin{pmatrix} \mathbb{1}_n + \sigma_n^z & \sigma_n^x - i\sigma_n^y \\ \sigma_n^x + i\sigma_n^y & \mathbb{1}_n - \sigma_n^z \end{pmatrix}_a = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

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

$$(40) \quad \mathcal{R} : \mathbb{V}_{a_1} \otimes \mathbb{V}_{a_2} \otimes \mathbb{V}_{a_1} \otimes \mathbb{V}_{a_2}$$

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

which⁵ 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 found⁶, consider

⁵Note that \mathcal{R} can be rewritten as a 2×2 -matrix in either of the auxiliary spaces, as follows

$$(41) \quad \mathcal{R}_{a_1, a_2} = \begin{pmatrix} \left(u + \frac{i}{2}\right) \mathbb{1}_{a_2} + i \mathbf{S}_{a_2}^z & i \mathbf{S}_{a_2}^- \\ i \mathbf{S}_{a_2}^+ & \left(u + \frac{i}{2}\right) \mathbb{1}_{a_2} - i \mathbf{S}_{a_2}^z \end{pmatrix}_{a_1} = \begin{pmatrix} \left(u + \frac{i}{2}\right) \mathbb{1}_{a_1} + i \mathbf{S}_{a_1}^z & i \mathbf{S}_{a_1}^- \\ i \mathbf{S}_{a_1}^+ & \left(u + \frac{i}{2}\right) \mathbb{1}_{a_1} - i \mathbf{S}_{a_1}^z \end{pmatrix}_{a_2}.$$

⁶Note that the following permutation properties hold

$$(42) \quad \mathcal{P}_{n, a_1} \mathcal{P}_{n, a_2} = \mathcal{P}_{a_1, a_2} \mathcal{P}_{n, a_1} = \mathcal{P}_{n, a_2} \mathcal{P}_{a_2, a_1} \text{ and } \mathcal{P}_{a, b} = \mathcal{P}_{b, a}$$

[illegible]

which proves that

$$(43) \quad \mathcal{R}_{a_1, a_2}(u_1 - u_2) \mathcal{L}_{n, a_1}(u_1) \mathcal{L}_{n, a_2}(u_2) = \mathcal{L}_{n, a_1}(u_2) \mathcal{L}_{n, a_1}(u_1) \mathcal{R}_{a_1, a_2}(u_1 - u_2).$$

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

$$(44) \quad \begin{aligned} \mathcal{L}_{n, 1} \mathcal{L}_{n, 2} \mathcal{L}_{n, 3} &= \mathcal{R}_{12}^{-1} \mathcal{L}_{n, 2} \mathcal{L}_{n, 1} \mathcal{L}_{n, 3} \mathcal{R}_{12} \\ &= \mathcal{R}_{12}^{-1} \mathcal{R}_{13}^{-1} \mathcal{L}_{n, 2} \mathcal{L}_{n, 3} \mathcal{L}_{n, 1} \mathcal{R}_{13} \mathcal{R}_{12} \\ &= \mathcal{R}_{12}^{-1} \mathcal{R}_{13}^{-1} \mathcal{R}_{23}^{-1} \mathcal{L}_{n, 3} \mathcal{L}_{n, 2} \mathcal{L}_{n, 1} \mathcal{R}_{23} \mathcal{R}_{13} \mathcal{R}_{12} \\ &= (\mathcal{R}_{23} \mathcal{R}_{13} \mathcal{R}_{12})^{-1} \mathcal{L}_{n, 3} \mathcal{L}_{n, 2} \mathcal{L}_{n, 1} \mathcal{R}_{23} \mathcal{R}_{13} \mathcal{R}_{12}, \end{aligned}$$

and similarly

$$(45) \quad \begin{aligned} \mathcal{L}_{n, 1} \mathcal{L}_{n, 2} \mathcal{L}_{n, 3} &= \mathcal{R}_{23}^{-1} \mathcal{L}_{n, 1} \mathcal{L}_{n, 3} \mathcal{L}_{n, 2} \mathcal{R}_{23} \\ &= \mathcal{R}_{23}^{-1} \mathcal{R}_{13}^{-1} \mathcal{L}_{n, 3} \mathcal{L}_{n, 1} \mathcal{L}_{n, 2} \mathcal{R}_{13} \mathcal{R}_{23} \\ &= \mathcal{R}_{23}^{-1} \mathcal{R}_{13}^{-1} \mathcal{R}_{12}^{-1} \mathcal{L}_{n, 3} \mathcal{L}_{n, 2} \mathcal{L}_{n, 1} \mathcal{R}_{12} \mathcal{R}_{13} \mathcal{R}_{23} \\ &= (\mathcal{R}_{12} \mathcal{R}_{13} \mathcal{R}_{23})^{-1} \mathcal{L}_{n, 3} \mathcal{L}_{n, 2} \mathcal{L}_{n, 1} \mathcal{R}_{12} \mathcal{R}_{13} \mathcal{R}_{23}, \end{aligned}$$

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

$$(46) \quad \mathcal{R}_{12} \mathcal{R}_{13} \mathcal{R}_{23} = \mathcal{R}_{23} \mathcal{R}_{13} \mathcal{R}_{12},$$

which is the quantum Yang-Baxter equation.

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

$$(47) \quad \begin{aligned} \mathcal{T}_{N, a} &= \prod_{j=1}^{j=N} \mathcal{L}_{j, a} \\ \mathcal{T}_{N, a} &= \prod_{j=1}^{j=N} \begin{pmatrix} u \mathbb{1}_j + i \mathbf{S}_j^z & i \mathbf{S}_j^- \\ i \mathbf{S}_j^+ & u \mathbb{1}_j - i \mathbf{S}_j^z \end{pmatrix} \equiv \begin{pmatrix} A(u) & B(u) \\ C(u) & D(u) \end{pmatrix}, \end{aligned}$$

This matrix obeys the fundamental relation as well⁷, ie.

$$(49) \quad \mathcal{R}_{a_1, a_2}(u_1 - u_2) \mathcal{T}_{N, a_1}(u_1) \mathcal{T}_{N, a_2}(u_2) = \mathcal{T}_{N, a_2}(u_2) \mathcal{T}_{N, a_1}(u_1) \mathcal{R}_{a_1, a_2}(u_1 - u_2).$$

⁷For example, consider $N = 2$, then

$$(48) \quad \begin{aligned} \mathcal{R}_{a_1, a_2}(u_1 - u_2) \mathcal{T}_{N, a_1}(u_1) \mathcal{T}_{N, a_2}(u_2) &= \mathcal{R}_{a_1, a_2}(u_1 - u_2) \mathcal{T}_{2, a_1}(u_1) \mathcal{T}_{2, a_2}(u_2) \\ &= \mathcal{R}_{a_1, a_2}(u_1 - u_2) \mathcal{L}_{2, a_1}(u_1) \mathcal{L}_{2, a_1}(u_1) \mathcal{L}_{2, a_2}(u_2) \mathcal{L}_{1, a_2}(u_2) \\ &= \left(\mathcal{R}_{a_1, a_2}(u_1 - u_2) \mathcal{L}_{2, a_1}(u_1) \mathcal{L}_{2, a_2}(u_2) \right) \mathcal{L}_{1, a_1}(u_1) \mathcal{L}_{1, a_2}(u_2) \\ &= \mathcal{L}_{2, a_2}(u_2) \mathcal{L}_{2, a_1}(u_1) \left(\mathcal{R}_{a_1, a_2}(u_1 - u_2) \mathcal{L}_{1, a_1}(u_1) \mathcal{L}_{1, a_2}(u_2) \right) \\ &= \mathcal{L}_{2, a_2}(u_2) \mathcal{L}_{2, a_1}(u_1) \mathcal{L}_{1, a_2}(u_2) \mathcal{L}_{1, a_1}(u_1) \mathcal{R}_{a_1, a_2}(u_1 - u_2) \\ &= \mathcal{L}_{2, a_2}(u_2) \mathcal{L}_{1, a_2}(u_2) \mathcal{L}_{2, a_1}(u_1) \mathcal{L}_{1, a_1}(u_1) \mathcal{R}_{a_1, a_2}(u_1 - u_2) \\ &= \mathcal{T}_{n, a_1}(u_2) \mathcal{T}_{n, a_1}(u_1) \mathcal{R}_{a_1, a_2}(u_1 - u_2), \end{aligned}$$

where the fundamental commutation relation given by (43) was used. Note as well that \mathcal{L} -operators acting on different spaces commute as well, eg. $[\mathcal{L}_{n_1, a_1}, \mathcal{L}_{n_2, a_2}] \propto \delta_{n_1, n_2}$. The general fundamental relation can then be proved via mathematical induction for the general case.

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

$$(50) \quad \mathfrak{t}(u) = \text{Tr}_{\mathbb{V}_a} \mathcal{T}_{N,a}(u) = \text{Tr}_{\mathbb{V}_a} \begin{pmatrix} A(u) & B(u) \\ C(u) & D(u) \end{pmatrix} = A(u) + D(u),$$

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 \(49\)](#)'s double trace can then be written in terms of the transfer matrix, as follows

$$\begin{aligned} & \text{Tr}_{\mathbb{V}_{a_1}} \text{Tr}_{\mathbb{V}_{a_2}} \left[\mathcal{R}_{a_1,a_2}(u_1 - u_2) \mathcal{T}_{N,a_1}(u_1) \mathcal{T}_{N,a_2}(u_2) \right] \\ & \Rightarrow \text{Tr}_{\mathbb{V}_{a_1}} \text{Tr}_{\mathbb{V}_{a_2}} \left[\mathcal{R}_{a_1,a_2}(u_1 - u_2) \mathcal{T}_{N,a_1}(u_1) \mathcal{T}_{N,a_2}(u_2) - \mathcal{T}_{N,a_1}(u_2) \mathcal{T}_{N,a_1}(u_1) \mathcal{R}_{a_1,a_2}(u_1 - u_2) \right] = 0 \\ & \Rightarrow \text{Tr}_{\mathbb{V}_{a_1}} \text{Tr}_{\mathbb{V}_{a_2}} \left[\mathcal{R}_{a_1,a_2}(u_1 - u_2) \left(\mathcal{T}_{N,a_1}(u_1) \mathcal{T}_{N,a_2}(u_2) - \mathcal{T}_{N,a_2}(u_2) \mathcal{T}_{N,a_1}(u_1) \right) \right] = 0 \end{aligned}$$

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.

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)

$$\begin{aligned} (51) \quad \mathcal{T}_{N,a}(z_0) &= \prod_1^{j=N} \left[\left(\cancel{z_0} \rightarrow \frac{i}{2} \right) \mathbb{1}_{j,a} + i \mathcal{P}_{j,a} \right] = i^N \mathcal{P}_{N,a} \mathcal{P}_{N-1,a} \cdots \mathcal{P}_{1,a} = \\ &= i^N \mathcal{P}_{1,2} \mathcal{P}_{2,3} \cdots \mathcal{P}_{N-1,N} \mathcal{P}_{N,a} \\ &= i^N \prod_{j=1}^{N-1} \mathcal{P}_{j,j+1} \mathcal{P}_{N,a} \end{aligned}$$

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

$$\begin{aligned} (52) \quad \mathfrak{t}(z_0) &= \text{Tr}_{\mathbb{V}_a} \mathcal{T}_{N,a}(z_0) = \text{Tr}_{\mathbb{V}_a} i^N \mathcal{P}_{1,2} i^N \prod_{j=1}^{N-1} \mathcal{P}_{j,j+1} \mathcal{P}_{N,a} \\ &= i^N \prod_{j=1}^{N-1} \mathcal{P}_{j,j+1} \text{Tr}_{\mathbb{V}_a} \mathcal{P}_{N,a} = i^N \prod_{j=1}^{N-1} \mathcal{P}_{j,j+1} \text{Tr}_{\mathbb{V}_a} \frac{1}{2} \begin{pmatrix} \mathbb{1}_n + \sigma_n^z & \sigma_n^x - i\sigma_n^y \\ \sigma_n^x + i\sigma_n^y & \mathbb{1}_n - \sigma_n^z \end{pmatrix}_a \\ &= i^N \prod_{j=1}^{N-1} \mathcal{P}_{j,j+1} \mathbb{1}_N. \end{aligned}$$

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

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

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

$$(53) \quad \mathcal{U} = e^{iP}$$

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

$$\begin{aligned}
\frac{d\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} \\
&\quad + \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} \\
&\quad \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} \\
&\quad + \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} \\
&\quad \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} \\
&\quad + \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} \left[\prod_1^{j=N-1} \mathcal{P}_j + \prod_N^{j=N} \mathcal{P}_j \times \prod_1^{j=N-2} \mathcal{P}_j + \cdots + \prod_k^{j=N} \mathcal{P}_j \times \prod_1^{k+2} \mathcal{P}_j \right. \\
&\quad \left. + \cdots + \prod_3^{j=N} \mathcal{P}_j \times \prod_1^{j=1} \mathcal{P}_j + \prod_2^{j=N} \mathcal{P}_j \right] \\
&= i^{N-1} \sum_{n \in \mathbb{N}} \mathcal{P}_{N,a} \cdots \mathcal{P}_{n-1,a} \mathcal{P}_{n+1,a} \cdots \mathcal{P}_{1,a} \\
&= i^{N-1} \sum_{n \in \mathbb{N}} \mathcal{P}_{1,2} \mathcal{P}_{2,3} \cdots \mathcal{P}_{n-1,n+1} \cdots \mathcal{P}_{N-1,N} \mathcal{P}_{N,a} \\
&= i^{N-1} \sum_{n \in \mathbb{N}} \prod_{j=1}^{n-2} \mathcal{P}_{j,j+1} \times \mathcal{P}_{n-1,n+1} \times \prod_{j=n+2}^{N-1} \mathcal{P}_{j,j+1} \mathcal{P}_{N,a}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
(54) \quad \frac{d\mathbf{t}(u)}{du} \Big|_{u=z_0} &= \frac{d \text{Tr}_{\mathbb{V}_a} \mathcal{T}_{N,a}(u)}{du} \Big|_{u=z_0} \\
&= i^{N-1} \sum_{n \in \mathbb{N}} \prod_{j=1}^{n-2} \mathcal{P}_{j,j+1} \times \mathcal{P}_{n-1,n+1} \times \prod_{j=n+2}^{N-1} \mathcal{P}_{j,j+1} \times \text{Tr}_{\mathbb{V}_a} \mathcal{P}_{N,a} \\
&= i^{N-1} \sum_{n \in \mathbb{N}} \prod_{j=1}^{n-2} \mathcal{P}_{j,j+1} \times \mathcal{P}_{n-1,n+1} \times \prod_{j=n+2}^{N-1} \mathcal{P}_{j,j+1}.
\end{aligned}$$

Now, note that

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

Note that (31) can be rewritten in terms of the permutation operator as

$$(56) \quad \mathbf{H} = \frac{J}{2} \sum_n \left(\mathcal{P}_{n,n+1} - \mathbb{1}^{\otimes N} \right),$$

which in turn can be rewritten as

$$(57) \quad \mathbf{H} = \frac{J}{2} \left(i \left. \frac{d}{du} \log(\mathbf{t}(u)) \right|_{u=z_0} - N \mathbb{1}^{\otimes N} \right),$$

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

Diagonalizing the Hamiltonian. The only task left is to diagonalize the Hamiltonian, by diagonalizing the transfer matrix. The monodromy matrix $\mathbf{t}_{N,a}$ 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 \mathcal{T} be the monodromy matrix for the Heisenberg model. Then, $\forall u_1, u_2 \in \mathbb{C}$ then*

- $[\mathcal{T}_{ij}(u_1), \mathcal{T}_{ij}(u_2)] = 0$, $i, j = 1, 2$, 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

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

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

$$\begin{aligned}
(58) \quad \mathcal{T}_{N,a_1}(u_1) &= \begin{pmatrix} A(u_1) & B(u_1) \\ C(u_1) & D(u_1) \end{pmatrix} \otimes \mathbb{1}_{V_2}, \quad \mathcal{T}_{N,a_2}(u_2) = \mathbb{1}_{V_1} \otimes \begin{pmatrix} A(u_2) & B(u_2) \\ C(u_2) & D(u_2) \end{pmatrix} \\
&= \begin{pmatrix} A(u_1) & 0 & B(u_1) & 0 \\ 0 & A(u_1) & 0 & B(u_1) \\ C(u_1) & 0 & D(u_1) & 0 \\ 0 & C(u_1) & 0 & D(u_1) \end{pmatrix}, \quad \begin{pmatrix} A(u_2) & B(u_2) & 0 & 0 \\ C(u_2) & D(u_2) & 0 & 0 \\ 0 & 0 & A(u_2) & B(u_2) \\ 0 & 0 & C(u_2) & D(u_2) \end{pmatrix}
\end{aligned}$$

Recall that the \mathcal{R} -matrix is given by (40), which can be rewritten as

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

Then the fundamental commutation relation (49) establishes a relationship between the monodromy \mathcal{T} -matrices and the \mathcal{R} -matrix. Then, plugging in the previous results yields

$$\begin{aligned}
& \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} = \mathbf{0},
\end{aligned}$$

from which, it's been proven that $[\mathcal{T}_{ij}(u_1), \mathcal{T}_{ij}(u_2)] = 0$, $i, j = 1, 2$ 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 $u_1 \rightarrow u_2$ implies $u_1 - u_2 \rightarrow -(u_1 - u_2)$. 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.

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, $\omega_+^i = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}_i^\dagger$ describes i -th spin's state, with up spin. Then,

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

Therefore, consider the \mathcal{L} -matrix, defined by (38), its action on a tensor product of states can be written out as

$$(59) \quad \mathcal{L}_{N,a} v \otimes \omega_+^i \begin{pmatrix} \lambda + \frac{i}{2} & \cdots \\ 0 & \lambda - \frac{i}{2} \end{pmatrix} v \otimes \omega_+^i, \quad \forall v \in \mathbb{V}_a,$$

since \mathbf{S}^+ annihilates ω_+^i . Thus, the ground state 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.

$$(60) \quad \Omega^{\otimes N} = \bigotimes_{i=1}^N \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \text{ where } \begin{aligned} A(u)\Omega^{\otimes N} &= \left(u + \frac{i}{2}\right)^N \Omega^{\otimes N}, \\ D(u)\Omega^{\otimes N} &= \left(u - \frac{i}{2}\right)^N \Omega^{\otimes N}, \\ C(u)\Omega^{\otimes N} &= 0. \end{aligned}$$

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

$$(61) \quad |u_1, \dots, u_M\rangle = \prod_{j=1}^{M < N} B(u_j) \Omega^{\otimes N}, \text{ for some set } \{u_j\}_{j=1}^{M < N} \subset \mathbb{C}.$$

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 (61). Then,

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

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,

$$(62) \quad \begin{aligned} A(v) |u_1, \dots, u_M\rangle &= \left[f(v-u_1)B(u_1)A(v) + g(v-u_1)B(v)A(u_1) \right] \prod_{j=2}^{M < N} B(u_j) \Omega^{\otimes N} \\ &= \left[\prod_{k=1}^{\ell} f(v-u_k) \right] \alpha^N(v) |u_1, \dots, u_M\rangle + \mathcal{O}(A \times D), \end{aligned}$$

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

(63)

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

Now, note that

$$\begin{aligned}
(A(v) + D(v)) |u_1, \dots, u_M\rangle &= \left[\left(\prod_{k=1}^{\ell} f(v - u_k) \right) \alpha^N(v) + \left(\prod_{k=1}^{\ell} h(v - u_k) \right) \delta^N(v) \right] |u_1, \dots, u_M\rangle \\
&\quad + \sum_{k=1}^{\ell} \left[\left(g(v - u_k) \left[\prod_{j \neq k}^{\ell} f(u_k - u_j) \right] \alpha^N(u_k) \right) + h(v - u_k) \left(\prod_{j \neq k}^{\ell} h(u_k - u_j) \right) \delta^N(u_j) \right. \\
&\quad \left. \times \prod_{\substack{j=1 \\ j \neq k}}^{\ell} \Omega^{\otimes N} B(u_i) \right] \\
&\equiv \left[\left(\prod_{k=1}^{\ell} f(v - u_k) \right) \alpha^N(v) + \left(\prod_{k=1}^{\ell} h(v - u_k) \right) \delta^N(v) \right] |u_1, \dots, u_M\rangle + 0 \\
&\Rightarrow \sum_{k=1}^{\ell} \left[\left(g(v - u_k) \left[\prod_{j \neq k}^{\ell} f(u_k - u_j) \right] \alpha^N(u_k) \right) + h(v - u_k) \left(\prod_{j \neq k}^{\ell} h(u_k - u_j) \right) \delta^N(u_j) \right] = 0 \\
&\quad \sum_{k=1}^{\ell} \left[\left(g(v - u_k) \left[\prod_{j \neq k}^{\ell} f(u_k - u_j) \right] \alpha^N(u_k) \right) - g(v - u_k) \left(\prod_{j \neq k}^{\ell} h(u_k - u_j) \right) \delta^N(u_j) \right] = 0 \\
&\quad \sum_{k=1}^{\ell} g(v - u_k) \left[\prod_{j \neq k}^{\ell} f(u_k - u_j) \right] \alpha^N(u_k) + \prod_{j \neq k}^{\ell} h(u_k - u_j) \delta^N(u_j) = 0,
\end{aligned}$$

which must be true for all k ,

$$(64) \quad \left[\prod_{j \neq k}^{\ell} f(v - u_k) \right] \alpha^N(u_j) = \left[\prod_{j \neq k}^{\ell} h(v - u_k) \right] \delta^N(u_j).$$

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

$$\begin{aligned}
&\left[\prod_{j \neq k}^{\ell} f(u_k - u_j) \right] \alpha^N(u_j) = \left[\prod_{j \neq k}^{\ell} h(u_k - u_j) \right] \delta^N(u_j) \\
(65) \quad &\prod_{j \neq k}^{\ell} \frac{u_k - u_j - i}{u_k - u_j} \left(u_j + \frac{i}{2} \right) = \prod_{j \neq k}^{\ell} \frac{u_k - u_j + i}{u_k - u_j} \left(u_j - \frac{i}{2} \right) \\
&0 = \prod_{j \neq k}^{\ell} \frac{u_k - u_j - i}{u_k - u_j} \left(u_j + \frac{i}{2} \right) - \prod_{j \neq k}^{\ell} \frac{u_k - u_j + i}{u_k - u_j} \left(u_j - \frac{i}{2} \right),
\end{aligned}$$

which can be rearranged to yield the Bethe ansatz equations,

$$(66) \quad \left(\frac{u_k + \frac{i}{2}}{u_k - \frac{i}{2}} \right)^N = \prod_{j \neq k}^{\ell} \frac{u_k - u_j + i}{u_k - u_j - i},$$

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,

$$(67) \quad \begin{aligned} (A(v) + D(v)) |u_1, \dots, u_M\rangle &= \left[\left(\prod_{k=1}^{\ell} f(v - u_k) \right) \alpha^N(v) + \left(\prod_{k=1}^{\ell} h(v - u_k) \right) \delta^N(v) \right] |u_1, \dots, u_M\rangle \\ (A(v) + D(v)) |u_1, \dots, u_M\rangle &= \left[\left(v + \frac{i}{2} \right)^N \prod_{k=1}^{\ell} \frac{v - u_k - i}{v - u_k} + \left(v - \frac{i}{2} \right)^N \prod_{k=1}^{\ell} \frac{v - u_k + i}{v - u_k} \right] |u_1, \dots, u_M\rangle \\ \left(A \left(\frac{i}{2} \right) + D \left(\frac{i}{2} \right) \right) |u_1, \dots, u_M\rangle &= i^N \prod_{k=1}^{\ell} \frac{u_k + \frac{i}{2}}{u_k - \frac{i}{2}} |u_1, \dots, u_M\rangle \equiv \lambda(\{u_k\}_{k=1}^{\ell < N}) |u_1, \dots, u_M\rangle \end{aligned}$$

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

$$(68) \quad P |u_1, \dots, u_{\ell}\rangle = \sum_{k=1}^{\ell} p(u_k) |u_1, \dots, u_{\ell}\rangle \text{ where } p(u_k) = \frac{1}{i} \log \frac{u_k + \frac{i}{2}}{u_k - \frac{i}{2}}.$$

Now, [equation \(57\)](#) 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,

$$\begin{aligned} \left. \frac{d \log \Lambda(v)}{dv} \right|_{v=z_0} &= \frac{1}{\Lambda(v)} \left. \frac{d \Lambda(v)}{dv} \right|_{v=z_0} \\ &= \frac{1}{\Lambda(v) \Big|_{v=z_0}} \left[N i^{N-1} \prod_{k=1}^{\ell} \frac{-u_k - \frac{i}{2}}{\frac{i}{2} - u_k} + i^N \sum_k \frac{i}{(\frac{i}{2} - u_k)^2} \prod_{\substack{j=1 \\ j \neq k}}^{\ell} \frac{-u_j - \frac{i}{2}}{\frac{i}{2} - u_j} \right] \\ &= i \left(\sum_{k=1}^{\ell} \frac{1}{u_k^2 + \frac{1}{4}} - N \right). \end{aligned}$$

Then, using [equation \(57\)](#), yields the eigenenergies

$$(69) \quad H |u_1, \dots, u_M\rangle = \left[\frac{i}{2} i \left(\sum_{k=1}^{\ell} \frac{1}{u_k^2 + \frac{1}{4}} - N \right) - \frac{N}{2} \right] |u_1, \dots, u_M\rangle$$

or, equivalently,

$$(70) \quad \mathbf{H} |u_1, \dots, u_M\rangle = \sum_{k=1}^{\ell} \epsilon(u_k) |u_1, \dots, u_M\rangle \text{ where } \epsilon(u_k) = -\frac{1}{2} \frac{1}{u_k^2 + \frac{1}{4}},$$

which concludes the Bethe solution for the XXX Heisenberg chain.

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

$$(71) \quad \mathbf{H} = J \sum_{j=1}^{L-1} (f_j^\dagger f_{j+1} + f_{j+1}^\dagger f_j) + \sum_{j=1} \lambda_j f_j^\dagger f_j, \quad \text{with the usual commutation rules} \quad \begin{cases} \{f_j, f_k\} = \{f_j^\dagger, f_k^\dagger\} = 0 \\ \{f_j, f_k^\dagger\} = \delta_{jk} \end{cases}$$

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 strength⁸. Said Hamiltonian has open boundaries conditions since there is no hopping term across the boundary. Note that we can rewrite (71) as

$$(72) \quad \mathbf{H} = \sum_{i,j=1}^L \mathcal{M}_{ij} f_i^\dagger f_j \quad \text{with} \quad \mathcal{M}_{ij} = \begin{cases} \lambda_i & \text{if } i = j \\ J & \text{if } j = i + 1 \text{ or } i = j + 1 \\ 0 & \text{otherwise} \end{cases}, \quad \mathcal{M} \in \text{GL}(L, \mathbb{R}),$$

which is a positive-defined tri-diagonal matrix. Let $\mathbf{f} = (f_1 \ f_2 \ \cdots \ f_L)^\text{T}$ be a vector of the L fermionic operators. Then, (71) can be rewritten as

$$(73) \quad \mathbf{H} = \mathbf{f}^\dagger \mathcal{M} \mathbf{f}.$$

Since \mathcal{M} is symmetric, then it can be diagonalized $\mathcal{M} = A \mathcal{D} A^\text{T}$, 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 (71) to be rewritten as

$$(74) \quad \mathbf{H} = \mathbf{f}^\dagger A \mathcal{D} A^\text{T} \mathbf{f} = \mathbf{d}^\dagger \mathcal{D} \mathbf{d}$$

where $\mathbf{d} = A^\text{T} \mathbf{f}$. Since the A -matrix is orthogonal, the new d_k -operators are fermionic operators as well, satisfying (71)'s anti-commutation rules. Then, the new fermionic operators are

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

Then, we can expand (74) in terms of the lattices, as follows

$$(75) \quad \mathbf{H} = \sum_{k=1}^L \epsilon_k d_k^\dagger d_k,$$

which is a sum of number operators with potentials. The eigenstates can then be constructed from the 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

$$(76) \quad \frac{d}{dt} d_k = i[\mathbf{H}, d_k],$$

and using that $d_k^2 = 0$, it turns out that (76)'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

⁸The 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.

$$(77) \quad \mathcal{N}_{jk} = \langle d_j^\dagger d_k \rangle = \begin{cases} 0 \text{ or } 1 & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases},$$

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,

$$(78) \quad \mathcal{N}_{jk} = \langle d_j^\dagger d_k \rangle_{\text{th}} = \frac{1}{1 + e^{\beta\epsilon_k + \mu}} \delta_{jk}.$$

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

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

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 (79), we can evolve any number operator or two-body correlation operator, ie.

$$(79) \quad \langle f_j^\dagger f_k \rangle = \sum_{k,l=1}^n A_{jk} A_{jl} \langle d_k^\dagger d_l \rangle.$$

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

$$(80) \quad \langle f_j^\dagger(t) f_k(t) \rangle = \sum_{k,l=1}^L e^{i(\epsilon_k - \epsilon_l)t} A_{jk} A_{jl} \langle d_k^\dagger d_l \rangle,$$

which can then be numerically solved.

2. 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.

Jellium oscillations and Einstein phonons. Let ρ_{ion}^0 be the particle density of the ion jellium and let $\rho_{\text{el}}^0 = Z\rho_{\text{ion}}^0$ that of the homogeneous electron gas. Consider slow ionic density oscillations in a static electron gas, where the restoring force is the long range Coulomb interaction, where the electron dynamics are neglected as well. This entails that ρ_{ext} is a fixed constant. In the limit of small harmonic deviations from the equilibrium, $\delta\rho_{\text{ion}}(x^\mu) = \delta\rho_{\text{ion}}(\mathbf{x})e^{-i\Omega t}$. The equations of motion are linear up to first order in the ionic density variations, thus its solutions can be written as

$$(81) \quad \rho_{\text{ion}}(x^\mu) = \rho_{\text{ion}}^0 + \delta\rho_{\text{ion}}(\mathbf{x})e^{-i\Omega t}.$$

A non-zero $\delta\rho_{\text{ion}}$ corresponds to a charge density $Ze\delta\rho_{\text{ion}}$ and hence is associated with an electric field \mathbf{E} which must obey Maxwell's equations, ie.

$$(82) \quad \nabla \cdot \mathbf{E} = \frac{Ze}{\epsilon_0} \delta\rho_{\text{ion}} \quad \Rightarrow \quad \nabla \cdot \mathbf{f} = \frac{Z^2 e^2 \rho_{\text{ion}}^0}{\epsilon_0} \delta\rho_{\text{ion}},$$

where $\mathbf{f} = Ze\rho_{\text{ion}}\mathbf{E} \approx Ze\rho_{\text{ion}}^0\mathbf{E}$ is a force density written upto first order in $\delta\rho_{\text{ion}}$. This force equation is supplemented by the continuity equation, $\partial_\mu J^\mu = 0$, where $J^\mu = (\rho_{\text{ion}}, \rho_{\text{ion}}\mathbf{v})$. This equation can be rewritten as

$$\partial_t \delta\rho_{\text{ion}} + \rho_{\text{ion}}^0 \nabla \cdot \mathbf{v} + \mathcal{O}(\delta^2 \rho_{\text{ion}}) = 0,$$

since the velocity is already a small quantity. Differentiating this with respect to time once and again and using Newton's second law $\mathbf{f} = M\rho_{\text{ion}}\partial_t \mathbf{v}$ yields

3. 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

$$(83) \quad \langle \mathbf{A} \rangle = \frac{1}{\mathcal{Z}_0} \text{Tr}_{\mathbb{H}} \rho_0 \mathbf{A} = \frac{1}{\mathcal{Z}_0} \sum_{n \in \Lambda} \langle n | \mathbf{A} | n \rangle e^{-\beta E_n} \quad \text{where } \{|n\rangle\}_{n \in \Lambda} \text{ is a complete set of eigenstates.}$$

$$\rho_0 = e^{-\beta \mathbf{H}_0} = \sum_{n \in \Lambda} |n\rangle \langle n| e^{-\beta E_n},$$

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

$$(84) \quad \mathbf{H}(t) = \mathbf{H}_0 + \mathbf{H}'(t)\theta(t - t_0)$$

Now, the interest lies in finding the expectation value of the \mathbf{A} operator at times t greater than 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

$$(85) \quad \langle \mathbf{A}(t) \rangle = \frac{1}{\mathcal{Z}_0} \text{Tr}_{\mathbb{H}} \rho(t) \mathbf{A} = \frac{1}{\mathcal{Z}_0} \sum_{n \in \Lambda} \langle n(t) | \mathbf{A} | n(t) \rangle e^{-\beta E_n}$$

$$(86) \quad \rho_0 = e^{-\beta \mathbf{H}_0} = \sum_{n \in \Lambda} |n(t)\rangle \langle n(t)| e^{-\beta E_n},$$

The physical idea behind this expression is as follows. The initial states of the system are distributed according to the usual Boltzmann distribution $\frac{e^{-\beta E_{0n}}}{\mathcal{Z}_0}$. 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

$$(87) \quad |n(t)\rangle = e^{-i\mathbf{H}_0 t} |n(t)\rangle_I = e^{-i\mathbf{H}_0 t} \mathcal{U}(t, t_0) |\hat{n}(t_0)\rangle,$$

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

$$\begin{aligned}
(88) \quad \langle \mathbf{A}(t) \rangle &= \frac{1}{\mathcal{Z}_0} \sum_{n \in \Lambda} \langle n(t) | \mathbf{A} | n(t) \rangle e^{-\beta E_n} \\
&= \frac{1}{\mathcal{Z}_0} \sum_{n \in \Lambda} \langle n(t_0) | \left(\mathbb{1} - i \int_{t_0}^t dt' \mathbf{H}'(t') \right) e^{-i\mathbf{H}_0 t_0} \mathbf{A} e^{i\mathbf{H}_0 t_0} \left(\mathbb{1} - i \int_{t_0}^t dt' \mathbf{H}'(t') \right) | n(t_0) \rangle e^{-\beta E_n} + \mathcal{O}(\mathbf{H}'^2) \\
&= \frac{1}{\mathcal{Z}_0} \sum_{n \in \Lambda} \langle n(t_0) | \mathbf{A} | n(t_0) \rangle - \frac{i}{\mathcal{Z}_0} \int_{t_0}^t dt' \sum_{n \in \Lambda} e^{-\beta E_n} \left(e^{-i\mathbf{H}_0 t_0} \mathbf{A} e^{i\mathbf{H}_0 t_0} \mathbf{H}'(t') - \mathbf{H}'(t') e^{-i\mathbf{H}_0 t_0} \mathbf{A} e^{i\mathbf{H}_0 t_0} | n(t_0) \rangle \right) \\
&= \langle \mathbf{A} \rangle_0 + \int_{t_0}^t \frac{dt'}{i} \langle [\mathbf{A}(t), \mathbf{H}'(t')] \rangle_0,
\end{aligned}$$

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.

$$(89) \quad \delta \langle \mathbf{A}(t) \rangle = \int_{t_0}^{\infty} dt' \mathcal{C}_{AH'}^R(t, t') e^{-\eta(t-t')} \text{ where } \mathcal{C}_{AH'}^R(t, t') = -i\theta(t-t') \langle [\mathbf{A}(t), \mathbf{H}'(t')] \rangle_0,$$

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 $t \gg t'$. At the end of the calculation, the limit $\eta \rightarrow 0^+$. This is so since the retarded effect of a perturbation must decrease in time⁹.

Kubo formula in the frequency domain. It is often convenient to express the response to an external disturbance in the frequency domain via Fourier transformations¹⁰. Therefore, consider the perturbation Hamiltonian \mathbf{H}' , which can be rewritten in terms of its Fourier components

$$(91) \quad \mathbf{H}'(t) = \int_{\Omega \subset \mathbb{R}} \frac{d\omega}{2\pi} e^{-i\omega t} \mathbf{H}'_{\omega},$$

such that the retarded correlation function becomes

$$(92) \quad \mathcal{C}_{AH'}^R(t, t') = \int_{\mathbb{R}} \frac{d\omega}{2\pi} e^{-i\omega t'} \mathcal{C}_{AH'_{\omega}}^R(t - t') \text{ since } \langle [\mathbf{A}(t), \mathbf{H}'(t')]_{\omega'} \rangle_0 \text{ only depends on the difference between } t \text{ and } t'.$$

Therefore, inserting this result into the Kubo formula yields

⁹The 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.

¹⁰Consider 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

$$(90) \quad \mathcal{F}[f(t)](\omega) = f(\omega) = \int_{\mathbb{R}} dt e^{i\omega t} f(t) \text{ and } \mathcal{F}^{-1}[f(\omega)](t) = f(t) = \int_{\mathbb{R}} \frac{d\omega}{2\pi} e^{-i\omega t} f(\omega)$$

$$\begin{aligned}
\delta\langle\mathbf{A}(t)\rangle &= \int_{t_0}^{\infty} dt' \mathcal{C}_{AH'_\omega}^R(t, t') e^{-\eta(t-t')} \\
&= \int_{\mathbb{R}} dt' \int_{\mathbb{R}} \frac{d\omega}{2\pi} e^{-i\omega t} e^{-i(\omega+i\eta)(t'-t)} \mathcal{C}_{AH'_\omega}^R(t-t') \\
(93) \quad &= \int_{\mathbb{R}} \frac{d\omega}{2\pi} e^{-i\omega t} \left(\int_{\mathbb{R}} d(t'-t) e^{-i(\omega+i\eta)(t'-t)} \mathcal{C}_{AH'_\omega}^R(t-t') \right) \\
&= \int_{\mathbb{R}} \frac{d\omega}{2\pi} e^{-i\omega t} \mathcal{C}_{AH'_\omega}^R(\omega; \eta)
\end{aligned}$$

which can be inverted to yield the final result in the frequency domain

$$\begin{aligned}
\int_{\mathbb{R}} dt e^{i\nu t} \delta\langle\mathbf{A}(t)\rangle &= \int_{\mathbb{R}} dt e^{i\nu t} \int_{\mathbb{R}} \frac{d\omega}{2\pi} e^{-i\omega t} \mathcal{C}_{AH'_\omega}^R(\omega; \eta) \\
(94) \quad \langle\mathbf{A}_\nu\rangle &= \int_{\mathbb{R}} \frac{d\omega}{2\pi} \int_{\mathbb{R}} dt e^{i(\nu-\omega)t} \mathcal{C}_{AH'_\omega}^R(\omega; \eta) \\
&= \int_{\mathbb{R}} d\omega \delta(\nu - \omega) \mathcal{C}_{AH'_\omega}^R(\omega; \eta) \\
&= \mathcal{C}_{AH'_\nu}^R(\nu; \eta)
\end{aligned}$$

$$(95) \quad \Rightarrow \delta\langle\mathbf{A}_\omega\rangle = \mathcal{C}_{AH'_\omega}^R(\omega) \text{ with } \mathcal{C}_{AH'_\omega}^R(\omega) = \int_{\mathbb{R}} dt e^{i\omega t} e^{-\eta t} \mathcal{C}_{AH'_\omega}^R(t),$$

where the infinitesimal η parameter is incorporated in order to ensure the correct physical result, ie. the retarded response function decays at $t \gg 1$.

Kubo formula for conductivity. Consider a system of charged particles, eg. electrons, which is subjected to 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.¹¹

$$(96) \quad J_e^\alpha(x^\mu) = \int_{\Omega \subset \mathbb{R}^4} dx^\mu \sum_{\alpha\beta} \sigma_\beta(x^\mu, y^\mu) E^\beta(y^\mu),$$

where $\sigma_{\alpha\beta}(x^\mu, y^\mu)$ is the conductivity tensor, which describes the current response in the \mathbf{e}_α -direction to an applied electric field in the \mathbf{e}_β -direction.

The electric field \mathbf{E} is given by the electric potential ϕ_{ext} and the vector potential \mathbf{A}_{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,

$$\mathbf{H}_{ext} = -e \int_{\mathbb{R}^3} d\mathbf{x} J_\mu(\mathbf{x}) A_{ext}^\mu(x^\mu).$$

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,

$$\mathbf{A}(x^\mu) = \mathbf{A}_0(x^\mu) + \mathbf{A}_{ext}(x^\mu).$$

The current operator can be decomposed in two components, the diamagnetic and the paramagnetic terms, as follows

$$(97) \quad \mathbf{J} = \mathbf{J}^\nabla(\mathbf{x}) + \frac{e}{m} \mathbf{A}(\mathbf{x}) \rho(\mathbf{x}).$$

¹¹Note that this section's mathematical treatment is not Lorentz-covariant. This is, the spacetime is treated as \mathbb{R}^4 with the metric being $g_{\mu\nu} = \delta_{\mu\nu}$

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

$$(98) \quad \partial_t \xrightarrow{F} -i\omega \quad \text{then} \quad \mathbf{A}_{\text{ext}}(\mathbf{x}, \omega) \xrightarrow{F} \frac{1}{i\omega} \mathbf{E}_{\text{ext}}(\mathbf{x}, \omega) \Rightarrow \mathbf{H}_{\text{ext}, \omega} = \frac{e}{i\omega} \int_{\Omega \subset \mathbb{R}^3} d\mathbf{x} \, \mathbf{J}(\mathbf{x}) \cdot \mathbf{E}_{\text{ext}}(\mathbf{x}, \omega).$$

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

$$(99) \quad J_e^\alpha(\mathbf{x}, \omega) = \int_{\Omega \subset \mathbb{R}^3} d\mathbf{y} \sum_{\beta} \sigma_{\alpha\beta}(\mathbf{x}, \mathbf{y}, \omega) E^\beta(\mathbf{y}, \omega).$$

Now, given that the external perturbation, written in frequency space, is already linear in the external potential \mathbf{E}_{ext} , and given that the interest lies only on terms proportional to said perturbation, the conductivity can be rewritten as

4. .