Tomás Pérez
Condensed Matter Theory - Lecture Notes
September 1, 2022

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-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.

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 $T_{|s\rangle}\mathcal{M}$,

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

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 $\langle s| : T_{|s\rangle}\mathcal{M} \to \mathbb{C}$.

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

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

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,

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

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. 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 the Ising model, an interacting many-body system, where the energies E_n depends on mutual properties of the degrees of freedom. This is a spin system where the spin is contrained to point in one of two directions, these directions are described by a variable $\sigma_i = 1$ for the up direction and $\sigma_i = -1$ for the down direction. For an N-site lattice, there are therefore 2^N different configurations in the model. The simplest interaction is to assign one energy unit if the neighbouring spins are the same, and another if neighbouring spins are different, that is

$$E(\{\sigma_i\}_{i=1}^N) = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j, \text{ where } \{\sigma_i\}_{i=1}^N \in \times_{i=1}^N T^*\mathcal{M}_i \tag{3}$$

and where the sum is performed on all nearest-neighbour sites i and j ie. this sum is not over all sites, but rather over all bonds. Note that on the square lattice, there are two bonds for every site. The parameter J is the coupling strength, if $J > 0$ this models describes a ferromagnetic coupling, while $J < 0$ describes an antiferromagnetic coupling. The partition function is then

$$\mathcal{Z} = \sum_{\sigma_i = \pm 1} e^{-\beta E(\{\sigma_i\})}, \tag{4}$$

where the sum is performed over all 2^N different configurations. From the partition function, the expectation values of physical quantities can be calculated. For example

- Internal energy $\langle E \rangle = \frac{1}{\mathcal{Z}} \sum_n E_n e^{-\beta E_n} = -\frac{\partial}{\partial \beta} \log \mathcal{Z}$.
- Specific heat $C = \frac{\partial}{\partial T} \langle E \rangle = k \frac{\partial}{\partial T} T^2 \frac{\partial \log \mathcal{Z}}{\partial T}$
- Two-point correlator in the Ising model,

$$\langle \sigma_a \sigma_b \rangle = \sum_{\sigma_i} \sigma_a \sigma_b \mathbb{P}(\sigma_i) = \frac{1}{\mathcal{Z}} \sum_{\sigma_i = \pm 1} \sigma_a \sigma_b e^{-\beta E(\{\sigma_i\})},$$

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

$$-\langle \sigma_a \rangle = -\sum_{\sigma_i = \pm 1} \sigma_i e^{-\beta E(\{\sigma_i\})} = \sum_{\sigma_i = \mp 1} s_a e^{-\beta E(\{s_i\})} \text{ where } s_i = -\sigma_i. \tag{5}$$

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.

1.1.1. *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 just two degrees of freedom, called s_1 and s_2. In addition, the Ising spins can only take two values $s_i \in \mathcal{M} = \{\pm 1\}$. To compute the partition function, $\mathcal{Z}$, the system's energy must be known for every configuration (s_1, s_2), which is assumed to be

$$E(\mathbf{s}) = -Js_1s_2 - B(s_1 + s_2). \tag{6}$$

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

$$\mathcal{Z} = \sum_{s_i \in \times_{i=1}^{2} \mathcal{M}} e^{Ks_1s_2 + h(s_1+s_2)} \text{ where } \begin{matrix} K = \beta J \\ h = \beta B. \end{matrix} \tag{7}$$

It is readily seen that

$$\mathcal{Z}(K, h) = 2\cosh(2h)e^{K} + 2e^{-K}. \tag{8}$$

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[1]. The thermal average can then be found to be

$$\begin{aligned} \langle M \rangle &= \frac{1}{\mathcal{Z}} \sum_{s_i \in \times_{i=1}^{2} \mathcal{M}} \frac{1}{2}(s_1 + s_2) e^{Ks_1s_2 + h(s_1+s_2)} = \frac{1}{2}\frac{1}{\mathcal{Z}}\frac{\partial \mathcal{Z}(K,h)}{\partial h} \\ &= \frac{1}{2}\frac{\partial \log \mathcal{Z}(K,h)}{\partial h}. \end{aligned} \tag{9}$$

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

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

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

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,

$$\mathcal{Z}(K,h) = \sum_{s_i\in\mathcal{M}^{\times 2}} e^{K s_1 s_2 + \mathbf{h}\cdot\mathbf{s}} = e^{-\beta F(K,\mathbf{h})} \tag{11}$$

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

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

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

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

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

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

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

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

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

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 interaction 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,

$$\chi = \frac{1}{2}\frac{\partial\langle M\rangle}{\partial h} = \frac{1}{4}\frac{\partial^2[-\beta F]}{\partial h^2} = \langle M^2\rangle - \langle M\rangle^2 \tag{17}$$

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.

Consider a lattice of $N+1$ equally spaced points, numbered by a natural index $i = 0, 1, \cdots, N$. 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

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

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

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

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

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

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

$$\text{If } N=3 \text{ then } \mathcal{Z}_3(\text{fixed}) = \textstyle\sum_{\sigma_2=\pm1} \mathbb{P}(\sigma_1\sigma_2)\mathbb{P}(\sigma_2\sigma_3).$$

$$\text{If } N=4 \text{ then } \mathcal{Z}_4(\text{fixed}) = \textstyle\sum_{\substack{\sigma_2=\pm1\\ \sigma_3=\pm1}} \mathbb{P}(\sigma_1\sigma_2)\mathbb{P}(\sigma_2\sigma_3)\mathbb{P}(\sigma_3 s\sigma_4).$$

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

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

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

$$\begin{aligned} \mathcal{Z}_N(\text{free}) &= \sum_{\substack{\sigma_1=\pm1\\ \sigma_N=\pm1}} \langle\sigma_1|\,\boldsymbol{\mathcal{T}}^{N-1}\,|\sigma_N\rangle = \begin{pmatrix} 1 & 0 \end{pmatrix}\boldsymbol{\mathcal{T}}^{N-1}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \end{pmatrix}\boldsymbol{\mathcal{T}}^{N-1}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &\quad + \begin{pmatrix} 0 & 1 \end{pmatrix}\boldsymbol{\mathcal{T}}^{N-1}\begin{pmatrix} 1 \\ 0 \end{pmatrix}.\begin{pmatrix} 0 & 1 \end{pmatrix}\boldsymbol{\mathcal{T}}^{N-1}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \end{pmatrix}\boldsymbol{\mathcal{T}}^{N-1}\begin{pmatrix} 1 \\ 1 \end{pmatrix}. \end{aligned} \tag{22}$$

- 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

$$\begin{aligned} \mathcal{Z}_N(\text{periodic}) &= \sum_{\sigma_1=\pm1} \langle\sigma_1|\,\boldsymbol{\mathcal{T}}\,|\sigma_1\rangle = \begin{pmatrix} 1 & 0 \end{pmatrix}\boldsymbol{\mathcal{T}}^{N}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \end{pmatrix}\boldsymbol{\mathcal{T}}^{N}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \operatorname{Tr}_{\mathbb{C}^2}\boldsymbol{\mathcal{T}}^{N}. \end{aligned} \tag{23}$$

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

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

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 $\boldsymbol{\mathcal{T}}$- 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,

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

- 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

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

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

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

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

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

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

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

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

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

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

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

then

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

Note that the answer factorizes over i since the Boltzmann weight factorizes over i as well, when written in terms of the $\mathfrak{t}_i$[2]. Then, the thermal average for any one $\mathfrak{t}$ can be easily calculated as

$$\begin{gathered}\langle\mathfrak{t}\rangle = \tfrac{1}{\mathcal{Z}}\textstyle\sum_{\mathfrak{t}=\pm1}\mathfrak{t}e^{E\{\mathfrak{t}\}} \\ \langle\mathfrak{t}\rangle = \tfrac{1e^{\beta J}-1e^{-\cdot\beta J}}{e^{\beta J}+e^{-\cdot\beta J}} = \tanh\beta J \quad\Rightarrow\quad \langle\sigma_j\sigma_i\rangle = (\tanh\beta J)^{|j-i|} = \exp\Big(|j-i|\log\tanh\beta J\Big).\end{gathered} \tag{33}$$

At any finite βJ, since $\tanh\beta J < 1$, $\langle\sigma_j\sigma_i\rangle \underset{|j-i|\to\infty}{\longrightarrow} 0$. 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

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

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

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

$$\begin{aligned}\mathcal{Z} &= \sum_{\sigma_i=\pm1}\exp\left(\sum_{i=0}^{N-1}\beta J\sigma_i\sigma_{i+1}\right) = \sum_{\mathfrak{t}_i=\pm1}\exp\left(\sum_{i=0}^{N-1}\beta J\mathfrak{t}_i\right) \\ &= \sum_{\mathfrak{t}_i=\pm1}\prod_{i=0}^{N-1}e^{\beta J\mathfrak{t}_i},\end{aligned} \tag{32}$$

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.

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 sinsle 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

$$\mathbf{H}_{\mathbb{XXX}} = -J \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j = -J \sum_{i=1}^{N} \Big(\frac{1}{2}(\sigma_i^+ \sigma_{i+1}^- + \sigma_i^- \sigma_{i+1}^+) + \frac{1}{4}\sigma_i^z \sigma_{i+1}^z \Big), \tag{35}$$

  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

$$\mathbf{H}_{\mathbb{XXZ}} = -\sum_{\langle jk \rangle} (J_\perp(\mathbf{S}_j^x \mathbf{S}_k^x + \mathbf{S}_j^y \mathbf{S}_k^y) + J_z \mathbf{S}_j^z \mathbf{S}_k^z) = -\sum_{j=1}^{N} \Big(J_\perp(\sigma_j^+ \sigma_{j+1}^- + \sigma_j^- \sigma_{j+1}^+) + \frac{\Delta}{2}\sigma_j^z \sigma_{j+1}^z \Big), \tag{36}$$

  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 $\mathbf{S}^a_j \to -\mathbf{S}^a_j$.

- 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,

$$\begin{aligned} \text{triplet} : |\uparrow\uparrow\rangle, \frac{1}{\sqrt{2}}\Big(|\downarrow\uparrow\rangle + |\uparrow\downarrow\rangle\Big), |\downarrow\downarrow\rangle, \text{ with } \lambda_i = -\frac{J}{4} \\ \text{singlet} : \frac{1}{\sqrt{2}}\Big(|\downarrow\uparrow\rangle - |\uparrow\downarrow\rangle\Big), \text{ with } \mu_i = +\frac{J}{4} \end{aligned} \tag{37}$$

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

$$\begin{array}{c} \mathcal{T} : \mathfrak{A}_i \to \mathfrak{A}_j \\ \mathcal{T}^{-1}\vec{\mathbf{S}}_i\mathcal{T} = \vec{\mathbf{S}}_{i+1} \end{array} \quad \text{where } \mathfrak{A}_i, \mathfrak{A}_j \simeq \mathfrak{su}(2)\ \forall i, j. \tag{38}$$

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

$$\begin{array}{c} |A\rangle = \frac{1}{\sqrt{2}}\Big(\uparrow\downarrow\uparrow\downarrow + e^{-i\pi k}\downarrow\uparrow\downarrow\uparrow\Big), \\ |B\rangle = \frac{1}{2}\Big(\left|\uparrow\uparrow\downarrow\downarrow\right\rangle + \left|\downarrow\uparrow\uparrow\downarrow\right\rangle + \cdots\Big) \end{array} \quad \begin{array}{c} \text{with the hamiltonian} \\ \text{on these} \\ k = 0 \text{ and } k = \pi \\ \text{states being} \end{array} \quad -J \begin{pmatrix} -1 & \sqrt{2}\cos\left(\frac{\pi k}{2}\right) \\ \sqrt{2}\cos\left(\frac{\pi k}{2}\right) & 0 \end{pmatrix}. \tag{39}$$

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

$$\begin{array}{l} \text{quintuplet} : 1 \\ \text{triplets} : \cos(\pi k)(k \neq 0) \\ \text{singlets} : -2, 0 \text{ for } k = 0, \pi \text{ respectively}. \end{array} \tag{40}$$

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.

$$\mathbf{H} = J\sum_{i=1}^{L}(\mathbf{S}_j^x\mathbf{S}_{j+1}^x + \mathbf{S}_j^y\mathbf{S}_{j+1}^y) - \lambda\sum_{j=1}^{L}\mathbf{S}_j^z, \tag{41}$$

which describes interacting spins in a one-dimensional chain, with periodic boundary conditions. (41)'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 (41) 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.

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

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

$$\mathbf{H} = \frac{J}{2}\sum_{i=1}^{L}(\mathbf{S}_j^+\mathbf{S}_{j+1}^- + \mathbf{S}_j^-\mathbf{S}_{j+1}^+) - \lambda\sum_{j=1}^{L}\mathbf{S}_j^z \tag{42}$$

where the interacting terms in the first summation, flip neighboring spins if said spins are anti-aligned[3]. In addition, the XX-Hamiltonian has a total magnetization symmetry, since the Hamiltonian given by (42) conmutes with the magnetization operator.

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

$$\begin{aligned}\mathbf{S}_j^z &= f_j^\dagger f_j - \tfrac{1}{2}\\ \mathbf{S}_j^- &= \exp\left(i\pi\textstyle\sum_{\ell=1}^{L-1} f_\ell^\dagger f_\ell\right)\\ \mathbf{S}_j^+ &= \exp\left(-i\pi\textstyle\sum_{\ell=1}^{L-1} f_\ell^\dagger f_\ell\right)\end{aligned} \tag{44}$$

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),

[3] 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.

$$(\mathbf{S}_1^+\mathbf{S}_2^- + \mathbf{S}_1^-\mathbf{S}_2^+)\,|\downarrow\downarrow\rangle = |\downarrow\downarrow\rangle \text{ and } (\mathbf{S}_1^+\mathbf{S}_2^- + \mathbf{S}_1^-\mathbf{S}_2^+)\,|\uparrow\uparrow\rangle = |\uparrow\uparrow\rangle \tag{43}$$

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

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

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

$$\mathbf{S}_L^+\mathbf{S}_1^- = \mathcal{Q} f_L^\dagger f_1 \text{ and } \mathbf{S}_L^-\mathbf{S}_1^+ = \mathcal{Q} f_1^\dagger f_L, \tag{45}$$

then (42) can be rewritten as

$$\mathbf{H} = \frac{J}{2}\sum_{i=1}^{L-1}\Big(f_j^\dagger f_{j+1} + f_{j+1}^\dagger f_j\Big) - \lambda\sum_{j=1}^{L}\Big(f_j^\dagger f_j - \frac{1}{2}\Big) + \frac{J}{2}\mathcal{Q}(f_L^\dagger f_1 + f_1^\dagger f_L), \tag{46}$$

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

$$\mathbf{H} = \frac{J}{2}\sum_{i=1}^{L}\Big(f_j^\dagger f_{j+1} + f_{j+1}^\dagger f_j\Big) - \lambda\sum_{j=1}^{L}\Big(f_j^\dagger f_j - \frac{1}{2}\Big) + \frac{J}{2}(\mathcal{Q}-1)(f_L^\dagger f_1 + f_1^\dagger f_L), \tag{47}$$

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 $\mathcal{O}\Big(\frac{1}{L}\Big)$-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

$$\mathbf{H} = \frac{J}{2}\sum_{i=1}^{L}\Big(f_j^\dagger f_{j+1} + f_{j+1}^\dagger f_j - \lambda f_j^\dagger f_j\Big) + \frac{\lambda L}{2}, \tag{48}$$

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

$$f_j = \frac{1}{\sqrt{L}}\sum_{\substack{k=2\pi m/L \\ m\in\mathbb{Z}_{[1,L]}}} e^{ijk} d_k, \qquad f_j^\dagger = \frac{1}{\sqrt{L}}\sum_{\substack{k=2\pi m/L \\ m\in\mathbb{Z}_{[1,L]}}} e^{-ijk} d_k^\dagger, \tag{49}$$

with the inverse transformation given by

$$d_k = \frac{1}{\sqrt{L}}\sum_{j=1}^{L} e^{-ikj} f_j \qquad d_k^\dagger = \frac{1}{\sqrt{L}}\sum_{j=1}^{L} e^{ikj} f_j^\dagger. \tag{50}$$

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, (48)'s terms are mapped as follows

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

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

$$\begin{aligned} \mathbf{H} &= \frac{J}{2} \sum_k \left(e^{ik} d_k^\dagger d_k + e^{-ik} d_k^\dagger d_k - \lambda d_k^\dagger d_k \right) + \frac{\lambda L}{2} \\ &= \sum_k \left(J \cos k - \lambda \right) d_k^\dagger d_k + \frac{\lambda L}{2}, \end{aligned} \tag{51}$$

which can be rewritten as

$$\mathbf{H} = \sum_k \epsilon_k d_k^\dagger d_k + \frac{\lambda L}{2} \quad \text{with the eigenvalues being given by } \epsilon_k = J \cos k - \lambda + \tfrac{\lambda L}{2} \text{ and the eigenvectors being given by } |E_n\rangle = \textstyle\prod_n (d_k^\dagger)^n, \text{ with eigenvalue } E_n = \sum_n \epsilon_n \tag{52}$$

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.

$$\mathcal{N}_{jk} = \langle d_j^\dagger d_k \rangle_{\text{th}} = \frac{1}{1 + e^{\beta \epsilon_k + \mu}} \delta_{jk} \tag{53}$$

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

$$\langle E_G|\, d_k^\dagger d_k\, |E_G\rangle = 1 \text{ if } \epsilon_k \leq 0 \text{ and } \langle E_G|\, d_k^\dagger d_k\, |E_G\rangle = 0 \text{ if } \epsilon_k > 0. \tag{54}$$

Then the groundstate energy is given by

$$E_G = \sum_k \epsilon_k \langle E_G|\, d_k^\dagger d_k\, |E_G\rangle\,. \tag{55}$$

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,

$$\begin{gathered} \langle E_G|\, d_k^\dagger d_k\, |E_G\rangle = 1 \qquad\qquad \langle E_G|\, d_k^\dagger d_k\, |E_G\rangle = 0 \\ \text{if } k \in [-\pi, -k_F] \cup [k_F, \pi], \qquad \text{and} \qquad \text{if } k \in [-k_F, k_F]. \end{gathered} \tag{56}$$

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

$$\mathbf{M}^z = \sum_k \left(d_k^\dagger d_k - \frac{1}{2}\right). \tag{57}$$

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

$$\langle E_G|\, \mathbf{M}^z\, |E_G\rangle = \sum_{k=-\pi}^{-k_F} 1 + \sum_{k=k_F}^{\pi} 1 - \frac{1}{2}\sum_{k=-\pi}^{\pi} 1 = \frac{1}{2}\Big(\sum_{k=-\pi}^{-k_F} 1 + \sum_{k=k_F}^{\pi} 1 - \sum_{k=-k_F}^{k_F} 1\Big), \tag{58}$$

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 the chain gets progressively bigger, with more and more lattice sites, the summations in (58) can be approximated by Riemannian integrals, ie.

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

Using this trick, (58) can be rewritten as

$$\langle \mathbf{m}^z\rangle_G = \frac{1}{4\pi}\Big(\int_{-\pi}^{-k_F} dk + \int_{k_F}^{\pi} dk - \int_{-k_F}^{k_F} dk\Big) = \frac{1}{4\pi}\Big(-k_F + \pi + \pi - k_F - k_F + k_F\Big) \tag{59}$$

$$= \frac{1}{2} - \frac{k_F}{\pi}, \tag{60}$$

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

$$\langle \mathbf{m}^z \rangle_G = \begin{cases} \frac{1}{2} - \frac{1}{\pi}\arccos\left(\frac{\lambda}{J}\right) & \text{if } |\lambda| \leq J \\ -\frac{1}{2} & \text{if } \lambda < J \\ \frac{1}{2} & \text{if } \lambda > J \end{cases} \tag{61}$$

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

$$\chi = \left(\frac{\partial \langle \mathbf{m}^z \rangle_G}{\partial \lambda}\right) \Rightarrow \chi = \begin{cases} \frac{1}{J\pi\sqrt{1-\frac{\lambda^2}{J^2}}} \text{ if } |\lambda| \leq J \\ 0 \text{ if } \lambda < J \\ 0 \text{ if } \lambda > J \end{cases}, \tag{62}$$

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

$$\begin{aligned}\mathbf{H} &= J\sum_{n=1}^{N}\left(\vec{\mathbf{S}}_n\cdot\vec{\mathbf{S}}_{n+1}-\frac{1}{4}\mathbb{1}^{\otimes N}\right)\\ &= \frac{J}{4}\sum_{n=1}^{N}\left(\vec{\sigma}_n\cdot\vec{\sigma}_{n+1}-\frac{1}{4}\mathbb{1}^{\otimes N}\right).\end{aligned}\tag{63}$$

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 $\boldsymbol{\mathcal{R}}$-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 $\boldsymbol{\mathcal{T}}$-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

$$A=\begin{pmatrix}A_{\mathcal{I}} & A_{\mathcal{I}\mathfrak{J}}\\ A_{\mathfrak{J}\mathcal{I}} & A_{\mathfrak{J}}\end{pmatrix}\begin{matrix}\mathbb{C}^k\\ \mathbb{C}^{n-k}\end{matrix}\quad\text{where}\quad\begin{matrix}A_{\mathcal{I}}\in\mathrm{GL}(k,\mathbb{C}), & A_{\mathcal{I}\mathfrak{J}}\in\mathrm{GL}(k\to n-k,\mathbb{C}),\\ A_{\mathfrak{J}\mathcal{I}}\in\mathrm{GL}(n-k\to k,\mathbb{C}), & A_{\mathfrak{J}}\in\mathrm{GL}(n-k,\mathbb{C}),\end{matrix}\tag{64}$$

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. [4]

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

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

$$\mathbb{H}\simeq\bigotimes_{n=1}^{N}\mathbb{V}_n\text{ where }\mathbb{V}_n\simeq\mathbb{C}^2\qquad\begin{matrix}\text{To these individual Hilbert spaces, additional}\\ \text{auxiliary, non-physical, spaces }\mathbb{V}_a\simeq\mathbb{C}^2\text{, can be added.}\end{matrix}\tag{65}$$

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

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

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

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

$$\begin{aligned}&\boldsymbol{\mathcal{L}} : \mathbb{V}_n \otimes \mathbb{V}_a \to \mathbb{V}_n \otimes \mathbb{V}_a \\ &\boldsymbol{\mathcal{L}}_{n,a} = u(\mathbb{1}_n \otimes \mathbb{1}_a) + i\sum_\alpha \mathbf{S}_n^\alpha \otimes \sigma_a^\alpha,\ \forall n \leq N\end{aligned} \tag{66}$$

where $\mathbb{1}_n$ and $\mathbf{S}_n^\alpha$ act on $\mathbb{V}_n$ 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 (66) can be rewritten as

$$\begin{aligned}\boldsymbol{\mathcal{L}}_{n,a} &= u(\mathbb{1}_n \otimes \mathbb{1}_a) + i\sum_\alpha \mathbf{S}_n^\alpha \otimes \sigma_a^\alpha \\ &= \begin{pmatrix} u\mathbb{1}_n + i\mathbf{S}_n^z & i\mathbf{S}_n^x + \mathbf{S}_n^y \\ i\mathbf{S}_n^x - \mathbf{S}_n^y & u\mathbb{1}_n - i\mathbf{S}_n^z \end{pmatrix}_a = \begin{pmatrix} u\mathbb{1}_n + i\mathbf{S}_n^z & i\mathbf{S}_n^- \\ i\mathbf{S}_n^+ & u\mathbb{1}_n - i\mathbf{S}_n^z \end{pmatrix}_a,\end{aligned} \tag{67}$$

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*

$$\boldsymbol{\mathcal{P}}_{n,a} = \frac{1}{2}\Big(\mathbb{1}_n \otimes \mathbb{1}_a + \sum_\alpha \sigma_n^\alpha \otimes \sigma_n^\alpha\Big) = \frac{1}{2}\sigma^{n\mu}\sigma_\mu^\alpha \tag{68}$$

which [5] *acts on both the physical and auxiliary systems.*

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

$$\begin{aligned}\boldsymbol{\mathcal{L}}_{n,a} &= u(\mathbb{1}_n \otimes \mathbb{1}_a) + i\sum_\alpha \mathbf{S}_n^\alpha \otimes \sigma_a^\alpha = u(\mathbb{1}_n \otimes \mathbb{1}_a) + \frac{i}{2}\sum_\alpha \sigma_n^\alpha \otimes \sigma_a^\alpha \\ &= u(\mathbb{1}_n \otimes \mathbb{1}_a) - \frac{i}{2}(\mathbb{1}_n \otimes \mathbb{1}_a) + \frac{i}{2}(\mathbb{1}_n \otimes \mathbb{1}_a) + \frac{i}{2}\sum_\alpha \sigma_n^\alpha \otimes \sigma_a^\alpha \\ &= \Big(u - \frac{i}{2}\Big)(\mathbb{1}_n \otimes \mathbb{1}_a) + \frac{i}{2}\Big(\mathbb{1}_n \otimes \mathbb{1}_a + \sum_\alpha \sigma_n^\alpha \otimes \sigma_n^\alpha\Big) \\ &= \Big(u - \frac{i}{2}\Big)\mathbb{1}_{n,a} + i\boldsymbol{\mathcal{P}}_{n,a}.\end{aligned} \tag{70}$$

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

$$\begin{aligned}&\boldsymbol{\mathcal{L}}_{n,a_1}(u_1) :\ \mathbb{V}_n \otimes \mathbb{V}_{a_1} \to \mathbb{V}_n \otimes \mathbb{V}_{a_1} \\ &\boldsymbol{\mathcal{L}}_{n,a_2}(u_2) :\ \mathbb{V}_n \otimes \mathbb{V}_{a_2} \to \mathbb{V}_n \otimes \mathbb{V}_{a_2}\end{aligned}$$

both of which act on the physical quantum state and on two different auxiliary spaces as well[6].

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

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

$$\begin{aligned}\boldsymbol{\mathcal{P}}_{n,a} &= \frac{1}{2}\left[\begin{pmatrix} \mathbb{1}_n & 0 \\ 0 & \mathbb{1}_a \end{pmatrix} + \sigma_n^x \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}_a + \sigma_n^y \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}_a + \sigma_n^z \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}_a\right] \\ &= \begin{pmatrix} \mathbb{1}_n + \sigma_n^z & \sigma_n^x - i\sigma_n^y \\ \sigma_n^x + i\sigma_n^y & \mathbb{1}_n - \sigma_n^z \end{pmatrix}_a = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\end{aligned} \tag{69}$$

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

$$\begin{aligned} &\boldsymbol{\mathcal{R}}:\ \mathbb{V}_{a_1}\otimes\mathbb{V}_{a_2}\otimes\mathbb{V}_{a_1}\otimes\mathbb{V}_{a_2}\\ &\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1-u_2)=(u_1-u_2)\mathbb{1}_{a_1,a_2}+i\boldsymbol{\mathcal{P}}_{a_1,a_2}, \end{aligned} \tag{72}$$

which[7] *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 $\boldsymbol{\mathcal{R}}$-matrix can then be found[8], consider

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

$$\boldsymbol{\mathcal{R}}_{a1,a2}=\begin{pmatrix}\left(u+\frac{i}{2}\right)\mathbb{1}_{a_2}+i\mathbf{S}^z_{a_2} & i\mathbf{S}^-_{a_2}\\ i\mathbf{S}^+_{a_2} & \left(u+\frac{i}{2}\right)\mathbb{1}_{a_2}-i\mathbf{S}^z_{a_2}\end{pmatrix}_{a_1}=\begin{pmatrix}\left(u+\frac{i}{2}\right)\mathbb{1}_{a_1}+i\mathbf{S}^z_{a_1} & i\mathbf{S}^-_{a_1}\\ i\mathbf{S}^+_{a_1} & \left(u+\frac{i}{2}\right)\mathbb{1}_{a_1}-i\mathbf{S}^z_{a_2}\end{pmatrix}_{a_2}. \tag{73}$$

[8]Note that the following permutation properties hold

$$\boldsymbol{\mathcal{P}}_{n,a_1}\boldsymbol{\mathcal{P}}_{n,a_2}=\boldsymbol{\mathcal{P}}_{a_1,a_2}\boldsymbol{\mathcal{P}}_{n,a_1}=\boldsymbol{\mathcal{P}}_{n,a_2}\boldsymbol{\mathcal{P}}_{a_2,a_1}\text{ and }\boldsymbol{\mathcal{P}}_{a,b}=\boldsymbol{\mathcal{P}}_{b,a} \tag{74}$$

$$
\begin{aligned}
&\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)\\
&\qquad=\Big((u_1-u_2)\mathbb{1}_{a_1,a_2}+i\boldsymbol{\mathcal{P}}_{a_1,a_2}\Big)\Big(\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_1}+i\boldsymbol{\mathcal{P}}_{n,a_1}\Big)\Big(\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}+i\boldsymbol{\mathcal{P}}_{n,a_2}\Big)\\
&\qquad=\Big((u_1-u_2)\mathbb{1}_{a_1,a_2}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_1}+(u_1-u_2)\mathbb{1}_{a_1,a_2}i\boldsymbol{\mathcal{P}}_{n,a_1}+i\boldsymbol{\mathcal{P}}_{a_1,a_2}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_1}+i^2\boldsymbol{\mathcal{P}}_{a_1,a_2}\boldsymbol{\mathcal{P}}_{n,a_1}\Big)\\
&\qquad\times\Big(\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}+i\boldsymbol{\mathcal{P}}_{n,a_2}\Big)\\
&\qquad=(u_1-u_2)\mathbb{1}_{a_1,a_2}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_1}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}+(u_1-u_2)\mathbb{1}_{a_1,a_2}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_1}i\boldsymbol{\mathcal{P}}_{n,a_2}\\
&\qquad+(u_1-u_2)\mathbb{1}_{a_1,a_2}i\boldsymbol{\mathcal{P}}_{n,a_1}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}+(u_1-u_2)\mathbb{1}_{a_1,a_2}i\boldsymbol{\mathcal{P}}_{n,a_1}i\boldsymbol{\mathcal{P}}_{n,a_2}\\
&\qquad+i\boldsymbol{\mathcal{P}}_{a_1,a_2}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_1}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}+i\boldsymbol{\mathcal{P}}_{a_1,a_2}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_1}i\boldsymbol{\mathcal{P}}_{n,a_2}\\
&\qquad-\boldsymbol{\mathcal{P}}_{a_1,a_2}\boldsymbol{\mathcal{P}}_{n,a_1}\Big(u-\frac{i}{2}\Big)\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}\\
&\qquad=(u_1-u_2)\mathbb{1}_{a_1,a_2}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}\times\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_1}+(u_1-u_2)\mathbb{1}_{a_1,a_2}i\boldsymbol{\mathcal{P}}_{n,a_2}\times\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_1}\\
&\qquad+(u_1-u_2)\mathbb{1}_{a_1,a_2}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}\times i\boldsymbol{\mathcal{P}}_{n,a_1}+(u_1-u_2)\mathbb{1}_{a_1,a_2}i\boldsymbol{\mathcal{P}}_{n,a_2}\times i\boldsymbol{\mathcal{P}}_{n,a_1}\\
&\qquad+i\boldsymbol{\mathcal{P}}_{a_1,a_2}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_1}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}+i\boldsymbol{\mathcal{P}}_{a_1,a_2}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_1}i\boldsymbol{\mathcal{P}}_{n,a_2}\\
&\qquad-\boldsymbol{\mathcal{P}}_{a_1,a_2}\boldsymbol{\mathcal{P}}_{n,a_1}\Big(u-\frac{i}{2}\Big)\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}\\
&\qquad=\Big((u_1-u_2)\mathbb{1}_{a_1,a_2}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}+(u_1-u_2)\mathbb{1}_{a_1,a_2}i\boldsymbol{\mathcal{P}}_{n,a_2}\Big)\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_1}\\
&\qquad+\Big((u_1-u_2)\mathbb{1}_{a_1,a_2}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}+(u_1-u_2)\mathbb{1}_{a_1,a_2}i\boldsymbol{\mathcal{P}}_{n,a_2}\Big)\times i\boldsymbol{\mathcal{P}}_{n,a_1}\\
&\qquad+i\boldsymbol{\mathcal{P}}_{a_1,a_2}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_1}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}+i\boldsymbol{\mathcal{P}}_{a_1,a_2}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_1}i\boldsymbol{\mathcal{P}}_{n,a_2}\\
&\qquad-\boldsymbol{\mathcal{P}}_{a_1,a_2}\boldsymbol{\mathcal{P}}_{n,a_1}\Big(u-\frac{i}{2}\Big)\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}\\
&\qquad=\Big((u_1-u_2)\mathbb{1}_{a_1,a_2}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}+(u_1-u_2)\mathbb{1}_{a_1,a_2}i\boldsymbol{\mathcal{P}}_{n,a_2}\Big)\Big(\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}+i\boldsymbol{\mathcal{P}}_{n,a_1}\Big)\\
&\qquad+i\boldsymbol{\mathcal{P}}_{a_1,a_2}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_1}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}+i\boldsymbol{\mathcal{P}}_{a_1,a_2}\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_1}i\boldsymbol{\mathcal{P}}_{n,a_2}\\
&\qquad-\boldsymbol{\mathcal{P}}_{a_1,a_2}\boldsymbol{\mathcal{P}}_{n,a_1}\Big(u-\frac{i}{2}\Big)\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}
\end{aligned}
$$

$$
\begin{aligned}
&=\Big(\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}+i\boldsymbol{\mathcal{P}}_{n,a_2}\Big)\Big(\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}+i\boldsymbol{\mathcal{P}}_{n,a_1}\Big)(u_1-u_2)\mathbb{1}_{a_1,a_2}\\
&+\Big(\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}+i\boldsymbol{\mathcal{P}}_{n,a_2}\Big)\Big(\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}+i\boldsymbol{\mathcal{P}}_{n,a_1}\Big)i\boldsymbol{\mathcal{P}}_{n,a_2}\\
&=\Big(\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}+i\boldsymbol{\mathcal{P}}_{n,a_2}\Big)\Big(\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}+i\boldsymbol{\mathcal{P}}_{n,a_1}\Big)\Big(\Big(u-\frac{i}{2}\Big)\mathbb{1}_{n,a_2}+i\boldsymbol{\mathcal{P}}_{n,a_2}\Big),
\end{aligned}
$$

which proves that

$$\boxed{\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1-u_2)\boldsymbol{\mathcal{L}}_{n,a_1}(u_1)\boldsymbol{\mathcal{L}}_{n,a_2}(u_2)=\boldsymbol{\mathcal{L}}_{n,a_1}(u_2)\boldsymbol{\mathcal{L}}_{n,a_1}(u_1)\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1-u_2)}. \tag{75}$$

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,

$$\begin{aligned}\boldsymbol{\mathcal{L}}_{n,1}\boldsymbol{\mathcal{L}}_{n,2}\boldsymbol{\mathcal{L}}_{n,3}&=\boldsymbol{\mathcal{R}}_{12}^{-1}\boldsymbol{\mathcal{L}}_{n,2}\boldsymbol{\mathcal{L}}_{n,1}\boldsymbol{\mathcal{L}}_{n,3}\boldsymbol{\mathcal{R}}_{12}\\&=\boldsymbol{\mathcal{R}}_{12}^{-1}\boldsymbol{\mathcal{R}}_{13}^{-1}\boldsymbol{\mathcal{L}}_{n,2}\boldsymbol{\mathcal{L}}_{n,3}\boldsymbol{\mathcal{L}}_{n,1}\boldsymbol{\mathcal{R}}_{13}\boldsymbol{\mathcal{R}}_{12}\\&=\boldsymbol{\mathcal{R}}_{12}^{-1}\boldsymbol{\mathcal{R}}_{13}^{-1}\boldsymbol{\mathcal{R}}_{23}^{-1}\boldsymbol{\mathcal{L}}_{n,3}\boldsymbol{\mathcal{L}}_{n,2}\boldsymbol{\mathcal{L}}_{n,1}\boldsymbol{\mathcal{R}}_{23}\boldsymbol{\mathcal{R}}_{13}\boldsymbol{\mathcal{R}}_{12}\\&=(\boldsymbol{\mathcal{R}}_{23}\boldsymbol{\mathcal{R}}_{13}\boldsymbol{\mathcal{R}}_{12})^{-1}\boldsymbol{\mathcal{L}}_{n,3}\boldsymbol{\mathcal{L}}_{n,2}\boldsymbol{\mathcal{L}}_{n,1}\boldsymbol{\mathcal{R}}_{23}\boldsymbol{\mathcal{R}}_{13}\boldsymbol{\mathcal{R}}_{12},\end{aligned} \tag{76}$$

and similarly

$$\begin{aligned}\boldsymbol{\mathcal{L}}_{n,1}\boldsymbol{\mathcal{L}}_{n,2}\boldsymbol{\mathcal{L}}_{n,3}&=\boldsymbol{\mathcal{R}}_{23}^{-1}\boldsymbol{\mathcal{L}}_{n,1}\boldsymbol{\mathcal{L}}_{n,3}\boldsymbol{\mathcal{L}}_{n,2}\boldsymbol{\mathcal{R}}_{23}\\&=\boldsymbol{\mathcal{R}}_{23}^{-1}\boldsymbol{\mathcal{R}}_{13}^{-1}\boldsymbol{\mathcal{L}}_{n,3}\boldsymbol{\mathcal{L}}_{n,1}\boldsymbol{\mathcal{L}}_{n,2}\boldsymbol{\mathcal{R}}_{13}\boldsymbol{\mathcal{R}}_{23}\\&=\boldsymbol{\mathcal{R}}_{23}^{-1}\boldsymbol{\mathcal{R}}_{13}^{-1}\boldsymbol{\mathcal{R}}_{12}^{-1}\boldsymbol{\mathcal{L}}_{n,3}\boldsymbol{\mathcal{L}}_{n,2}\boldsymbol{\mathcal{L}}_{n,1}\boldsymbol{\mathcal{R}}_{12}\boldsymbol{\mathcal{R}}_{13}\boldsymbol{\mathcal{R}}_{23}\\&=(\boldsymbol{\mathcal{R}}_{12}\boldsymbol{\mathcal{R}}_{13}\boldsymbol{\mathcal{R}}_{23})^{-1}\boldsymbol{\mathcal{L}}_{n,3}\boldsymbol{\mathcal{L}}_{n,2}\boldsymbol{\mathcal{L}}_{n,1}\boldsymbol{\mathcal{R}}_{12}\boldsymbol{\mathcal{R}}_{13}\boldsymbol{\mathcal{R}}_{23},\end{aligned} \tag{77}$$

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

$$\boxed{\boldsymbol{\mathcal{R}}_{12}\boldsymbol{\mathcal{R}}_{13}\boldsymbol{\mathcal{R}}_{23}=\boldsymbol{\mathcal{R}}_{23}\boldsymbol{\mathcal{R}}_{13}\boldsymbol{\mathcal{R}}_{12}}, \tag{78}$$

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.*

$$\begin{aligned}\boldsymbol{\mathcal{T}}_{N,a}&=\prod_1^{j=N}\boldsymbol{\mathcal{L}}_{j,a}\\\boldsymbol{\mathcal{T}}_{N,a}&=\prod_1^{j=N}\begin{pmatrix}u\mathbb{1}_j+i\mathbf{S}_j^z & i\mathbf{S}_j^-\\ i\mathbf{S}_j^+ & u\mathbb{1}_j-i\mathbf{S}_j^z\end{pmatrix}\equiv\begin{pmatrix}A(u) & B(u)\\ C(u) & D(u)\end{pmatrix},\end{aligned} \tag{79}$$

This matrix obeys the fundamental relation as well[9], ie.

$$\boxed{\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1-u_2)\boldsymbol{\mathcal{T}}_{N,a_1}(u_1)\boldsymbol{\mathcal{T}}_{N,a_2}(u_2)=\boldsymbol{\mathcal{T}}_{N,a_2}(u_2)\boldsymbol{\mathcal{T}}_{N,a_1}(u_1)\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1-u_2)}. \tag{81}$$

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

$$\begin{aligned}\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1-u_2)\boldsymbol{\mathcal{T}}_{N,a_1}(u_1)\boldsymbol{\mathcal{T}}_{N,a_2}(u_2)&=\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1-u_2)\boldsymbol{\mathcal{T}}_{2,a_1}(u_1)\boldsymbol{\mathcal{T}}_{2,a_2}(u_2)\\&=\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1-u_2)\boldsymbol{\mathcal{L}}_{2,a_1}(u_1)\boldsymbol{\mathcal{L}}_{2,a_1}(u_1)\boldsymbol{\mathcal{L}}_{2,a_2}(u_2)\boldsymbol{\mathcal{L}}_{1,a_2}(u_2)\\&=\Big(\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1-u_2)\boldsymbol{\mathcal{L}}_{2,a_1}(u_1)\boldsymbol{\mathcal{L}}_{2,a_2}(u_2)\Big)\boldsymbol{\mathcal{L}}_{1,a_1}(u_1)\boldsymbol{\mathcal{L}}_{1,a_2}(u_2)\\&=\boldsymbol{\mathcal{L}}_{2,a_2}(u_2)\boldsymbol{\mathcal{L}}_{2,a_1}(u_1)\Big(\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1-u_2)\boldsymbol{\mathcal{L}}_{1,a_1}(u_1)\boldsymbol{\mathcal{L}}_{1,a_2}(u_2)\Big)\\&=\boldsymbol{\mathcal{L}}_{2,a_2}(u_2)\boldsymbol{\mathcal{L}}_{2,a_1}(u_1)\boldsymbol{\mathcal{L}}_{1,a_2}(u_2)\boldsymbol{\mathcal{L}}_{1,a_1}(u_1)\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1-u_2)\\&=\boldsymbol{\mathcal{L}}_{2,a_2}(u_2)\boldsymbol{\mathcal{L}}_{1,a_2}(u_2)\boldsymbol{\mathcal{L}}_{2,a_1}(u_1)\boldsymbol{\mathcal{L}}_{1,a_1}(u_1)\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1-u_2)\\&=\boldsymbol{\mathcal{T}}_{n,a_1}(u_2)\boldsymbol{\mathcal{T}}_{n,a_1}(u_1)\boldsymbol{\mathcal{R}}_{a_1,a_2}(u_1-u_2),\end{aligned} \tag{80}$$

where the fundamental commutation relation given by (75) 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.

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.*

$$\mathfrak{t}(u) = \mathrm{Tr}_{\mathbb{V}_a}\, \boldsymbol{\mathcal{T}}_{N,a}(u) = \mathrm{Tr}_{\mathbb{V}_a} \begin{pmatrix} A(u) & B(u) \\ C(u) & D(u) \end{pmatrix} = A(u) + D(u), \tag{82}$$

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 (81)'s double trace can then be written in terms of the transfer matrix, as follows

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

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)

$$\begin{aligned} \boldsymbol{\mathcal{T}}_{N,a}(z_0) = \prod_{1}^{j=N} \Big[\cancel{\Big(z_0 - \frac{i}{2}\Big)} \mathbb{1}_{j,a} + i \boldsymbol{\mathcal{P}}_{j,a} \Big] &= i^N \boldsymbol{\mathcal{P}}_{N,a} \boldsymbol{\mathcal{P}}_{N-1,a} \cdots \boldsymbol{\mathcal{P}}_{1,a} = \\ &= i^N \boldsymbol{\mathcal{P}}_{1,2} \boldsymbol{\mathcal{P}}_{2,3} \cdots \boldsymbol{\mathcal{P}}_{N-1,N} \boldsymbol{\mathcal{P}}_{N,a} \\ &= i^N \prod_{j=1}^{N-1} \boldsymbol{\mathcal{P}}_{j,\,j+1} \boldsymbol{\mathcal{P}}_{N,a} \end{aligned} \tag{83}$$

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

$$\begin{aligned} \mathfrak{t}(z_0) = \mathrm{Tr}_{\mathbb{V}_a}\, \boldsymbol{\mathcal{T}}_{N,a}(z_0) &= \mathrm{Tr}_{\mathbb{V}_a}\, i^N \boldsymbol{\mathcal{P}}_{1,2} i^N \prod_{j=1}^{N-1} \boldsymbol{\mathcal{P}}_{j,\,j+1} \boldsymbol{\mathcal{P}}_{N,a} \\ &= i^N \prod_{j=1}^{N-1} \boldsymbol{\mathcal{P}}_{j,\,j+1}\, \mathrm{Tr}_{\mathbb{V}_a}\, \boldsymbol{\mathcal{P}}_{N,a} = i^N \prod_{j=1}^{N-1} \boldsymbol{\mathcal{P}}_{j,\,j+1}\, \mathrm{Tr}_{\mathbb{V}_a}\, \frac{1}{2} \begin{pmatrix} \mathbb{1}_n + \sigma_n^z & \sigma_n^x - i\sigma_n^y \\ \sigma_n^x + i\sigma_n^y & \mathbb{1}_n - \sigma_n^z \end{pmatrix}_a . \\ &= i^N \prod_{j=1}^{N-1} \boldsymbol{\mathcal{P}}_{j,\,j+1} \mathbb{1}_N . \end{aligned} \tag{84}$$

Then $\mathcal{U} = i^{-N} \mathfrak{t}(z_0) = \prod_{j=1}^{N-1} \boldsymbol{\mathcal{P}}_{j,\,j+1}$ 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

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

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

$$\begin{aligned}
\frac{d\boldsymbol{\mathcal{T}}_{N,a}(u)}{du}\Big|_{u=z_0} &= \frac{d}{du}\prod_{1}^{j=N}\begin{pmatrix} u\mathbb{1}_j + i\mathbf{S}_j^z & i\mathbf{S}_j^- \\ i\mathbf{S}_j^+ & u\mathbb{1}_j - i\mathbf{S}_j^z \end{pmatrix}\Big|_{u=z_0} \\
&= \frac{d}{du}\left[\begin{pmatrix} u\mathbb{1}_N + i\mathbf{S}_N^z & i\mathbf{S}_N^- \\ i\mathbf{S}_N^+ & u\mathbb{1}_N - i\mathbf{S}_N^z \end{pmatrix}\right]\Big|_{u=z_0}\prod_{1}^{j=N-1}\begin{pmatrix} u\mathbb{1}_j + i\mathbf{S}_j^z & i\mathbf{S}_j^- \\ i\mathbf{S}_j^+ & u\mathbb{1}_j - i\mathbf{S}_j^z \end{pmatrix} \\
&+ \begin{pmatrix} u\mathbb{1}_N + i\mathbf{S}_N^z & i\mathbf{S}_N^- \\ i\mathbf{S}_N^+ & u\mathbb{1}_N - i\mathbf{S}_N^z \end{pmatrix}\frac{d}{du}\left[\begin{pmatrix} u\mathbb{1}_{N-1} + i\mathbf{S}_{N-1}^z & i\mathbf{S}_{N-1}^- \\ i\mathbf{S}_{N-1}^+ & u\mathbb{1}_{N-1} - i\mathbf{S}_{N-1}^z \end{pmatrix}\right]\Big|_{u=z_0} \\
&\times\prod_{1}^{j=N-2}\begin{pmatrix} u\mathbb{1}_j + i\mathbf{S}_j^z & i\mathbf{S}_j^- \\ i\mathbf{S}_j^+ & u\mathbb{1}_j - i\mathbf{S}_j^z \end{pmatrix} \\
&+ \cdots + \prod_{3}^{j=N}\begin{pmatrix} u\mathbb{1}_j + i\mathbf{S}_j^z & i\mathbf{S}_j^- \\ i\mathbf{S}_j^+ & u\mathbb{1}_j - i\mathbf{S}_j^z \end{pmatrix}\frac{d}{du}\left[\begin{pmatrix} u\mathbb{1}_2 + i\mathbf{S}_2^z & i\mathbf{S}_2^- \\ i\mathbf{S}_2^+ & u\mathbb{1}_2 - i\mathbf{S}_2^z \end{pmatrix}\right]\Big|_{u=z_0} \\
&\times\begin{pmatrix} u\mathbb{1}_1 + i\mathbf{S}_1^z & i\mathbf{S}_1^- \\ i\mathbf{S}_1^+ & u\mathbb{1}_1 - i\mathbf{S}_1^z \end{pmatrix} \\
&+ \prod_{2}^{j=N}\begin{pmatrix} u\mathbb{1}_j + i\mathbf{S}_j^z & i\mathbf{S}_j^- \\ i\mathbf{S}_j^+ & u\mathbb{1}_j - i\mathbf{S}_j^z \end{pmatrix}\frac{d}{du}\left[\begin{pmatrix} u\mathbb{1}_1 + i\mathbf{S}_1^z & i\mathbf{S}_1^- \\ i\mathbf{S}_1^+ & u\mathbb{1}_1 - i\mathbf{S}_1^z \end{pmatrix}\right]\Big|_{u=z_0} s
\end{aligned}$$

$$\begin{aligned}
&= i^{N-1}\Big[\prod_{1}^{j=N-1}\boldsymbol{\mathcal{P}}_j + \prod_{N}^{j=N}\boldsymbol{\mathcal{P}}_j \times \prod_{1}^{j=N-2}\boldsymbol{\mathcal{P}}_j + \cdots + \prod_{k}^{j=N}\boldsymbol{\mathcal{P}}_j \times \prod_{1}^{k+2}\boldsymbol{\mathcal{P}}_j \\
&+ \cdots + \prod_{3}^{j=N}\boldsymbol{\mathcal{P}}_j \times \prod_{1}^{j=1}\boldsymbol{\mathcal{P}}_j + \prod_{2}^{j=N}\boldsymbol{\mathcal{P}}_j\Big] \\
&= i^{N-1}\sum_{n\in\mathbb{N}}\boldsymbol{\mathcal{P}}_{N,a}\cdots\boldsymbol{\mathcal{P}}_{n-1,a}\boldsymbol{\mathcal{P}}_{n+1,a}\cdots\boldsymbol{\mathcal{P}}_{1,a} \\
&= i^{N-1}\sum_{n\in\mathbb{N}}\boldsymbol{\mathcal{P}}_{1,2}\boldsymbol{\mathcal{P}}_{2,3}\cdots\boldsymbol{\mathcal{P}}_{n-1,n+1}\cdots\boldsymbol{\mathcal{P}}_{N-1,N}\boldsymbol{\mathcal{P}}_{N,a} \\
&= i^{N-1}\sum_{n\in\mathbb{N}}\prod_{j=1}^{n-2}\boldsymbol{\mathcal{P}}_{j,\,j+1}\times\boldsymbol{\mathcal{P}}_{n-1,\,n+1}\times\prod_{j=n+2}^{N-1}\boldsymbol{\mathcal{P}}_{j,\,j+1}\boldsymbol{\mathcal{P}}_{N,a}.
\end{aligned}$$

Therefore,

(86)
$$\begin{aligned}
\frac{d\mathfrak{t}(u)}{du}\Big|_{u=z_0} &= \frac{d\,\mathrm{Tr}_{\mathbb{V}_a}\boldsymbol{\mathcal{T}}_{N,a}(u)}{du}\Big|_{u=z_0} \\
&= i^{N-1}\sum_{n\in\mathbb{N}}\prod_{j=1}^{n-2}\boldsymbol{\mathcal{P}}_{j,\,j+1}\times\boldsymbol{\mathcal{P}}_{n-1,\,n+1}\times\prod_{j=n+2}^{N-1}\boldsymbol{\mathcal{P}}_{j,\,j+1}\times\mathrm{Tr}_{\mathbb{V}_a}\boldsymbol{\mathcal{P}}_{N,a} \\
&= i^{N-1}\sum_{n\in\mathbb{N}}\prod_{j=1}^{n-2}\boldsymbol{\mathcal{P}}_{j,\,j+1}\times\boldsymbol{\mathcal{P}}_{n-1,\,n+1}\times\prod_{j=n+2}^{N-1}\boldsymbol{\mathcal{P}}_{j,\,j+1}.
\end{aligned}$$

Now, note that

$$
\begin{aligned}
\frac{d}{du}\log(\mathfrak{t}(u))\Big|_{u=z_0} &= \frac{d}{du}\log \mathrm{Tr}_{\mathbb{V}_a}\,\boldsymbol{\mathcal{T}}(u)\Big|_{u=z_0}\\
&= \frac{d}{du}\Big(\mathfrak{t}(u)\Big)\mathfrak{t}(u)^{-1}\Big|_{u=z_0}\\
&= i^{N-1}\sum_{n\in\mathbb{N}}\Big(\prod_{j=1}^{n-2}\boldsymbol{\mathcal{P}}_{j,\,j+1}\times\boldsymbol{\mathcal{P}}_{n-1,\,n+1}\times\prod_{j=n+2}^{N-1}\boldsymbol{\mathcal{P}}_{j,\,j+1}\Big)\times\Big(i^{N}\prod_{j=1}^{N-1}\boldsymbol{\mathcal{P}}_{j,\,j+1}\mathbb{1}_N\Big)^{-1}\\
&= \frac{1}{i}\sum_{n\in\mathbb{N}}\Big(\prod_{j=1}^{n-2}\boldsymbol{\mathcal{P}}_{j,\,j+1}\times\boldsymbol{\mathcal{P}}_{n-1,\,n+1}\times\prod_{j=n+2}^{N-1}\boldsymbol{\mathcal{P}}_{j,\,j+1}\Big)\times\prod_{j=1}^{N-1}\boldsymbol{\mathcal{P}}^{-1}_{N-1-j,N-j}\\
&= \frac{1}{i}\sum_{n\in\mathbb{N}}\Big(\prod_{j=1}^{n-2}\boldsymbol{\mathcal{P}}_{j,\,j+1}\times\boldsymbol{\mathcal{P}}_{n-1,\,n+1}\times\prod_{j=n+2}^{N-1}\boldsymbol{\mathcal{P}}_{j,\,j+1}\Big)\times\prod_{1}^{j=N-1}\boldsymbol{\mathcal{P}}_{j,j+1}\\
&= \frac{1}{i}\sum_{n\in\mathbb{N}}\Big(\cancel{\boldsymbol{\mathcal{P}}_{1,2}\boldsymbol{\mathcal{P}}_{2,3}}\cdots\boldsymbol{\mathcal{P}}_{n-1,\,n+1}\cdots\cancel{\boldsymbol{\mathcal{P}}_{N-1,N}\times\boldsymbol{\mathcal{P}}_{N-1,\,N}}\cdots\boldsymbol{\mathcal{P}}_{n,\,n+1}\cdots\cancel{\boldsymbol{\mathcal{P}}_{2,3}\boldsymbol{\mathcal{P}}_{1,2}}\Big)\\
&= \frac{1}{i}\sum_{n\in\mathbb{N}}\boldsymbol{\mathcal{P}}_{N-1,\,N}
\end{aligned}
\tag{87}
$$

Note that (63) can be rewritten in terms of the permutation operator as

$$
\mathbf{H} = \frac{J}{2}\sum_{n}\Big(\boldsymbol{\mathcal{P}}_{n,\,n+1} - \mathbb{1}^{\otimes N}\Big),
\tag{88}
$$

which in turn can be rewritten as

$$
\mathbf{H} = \frac{J}{2}\Big(i\frac{d}{du}\log(\mathfrak{t}(u))\Big|_{u=z_0} - N\mathbb{1}^{\otimes N}\Big),
\tag{89}
$$

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 $\boldsymbol{\mathcal{R}}$ and $\boldsymbol{\mathcal{T}}$-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

$$
\begin{aligned}
(90)\qquad \boldsymbol{\mathcal{T}}_{N,a_1}(u_1) &= \begin{pmatrix} A(u_1) & B(u_1) \\ C(u_1) & D(u_1) \end{pmatrix} \otimes \mathbb{1}_{\mathbb{V}_2}, \quad \boldsymbol{\mathcal{T}}_{N,a_2}(u_2) = \mathbb{1}_{\mathbb{V}_1} \otimes \begin{pmatrix} A(u_2) & B(u_2) \\ C(u_2) & D(u_2) \end{pmatrix} \\
&= \begin{pmatrix} A(u_1) & 0 & B(u_1) & 0 \\ 0 & A(u_1) & 0 & B(u_1) \\ C(u_1) & 0 & D(u_1) & 0 \\ 0 & C(u_1) & 0 & D(u_1) \end{pmatrix}, \quad \begin{pmatrix} A(u_2) & B(u_2) & 0 & 0 \\ C(u_2) & D(u_2) & 0 & 0 \\ 0 & 0 & A(u_2) & B(u_2) \\ 0 & 0 & C(u_2) & D(u_2) \end{pmatrix}
\end{aligned}
$$

Recall that the $\boldsymbol{\mathcal{R}}$-matrix is given by (72), which can be rewritten as

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

Then the fundamental commutation relation (81) establishes a relationship between the monodromy $\boldsymbol{\mathcal{T}}$-matrices and the $\boldsymbol{\mathcal{R}}$-matrix. Then, plugging in the previous results yields

$$
\begin{aligned}
&\begin{pmatrix} a_{12} & 0 & 0 & 0 \\ 0 & b_{12} & c_{12} & 0 \\ 0 & c_{12} & b_{12} & 0 \\ 0 & 0 & 0 & a_{12} \end{pmatrix} \begin{pmatrix} A(u_1)A(u_2) & A(u_1)B(u_2) & B(u_1)A(u_2) & B(u_1)B(u_2) \\ A(u_1)C(u_2) & A(u_1)D(u_2) & B(u_1)C(u_2) & B(u_1)D(u_2) \\ C(u_1)A(u_2) & C(u_1)B(u_2) & D(u_1)A(u_2) & D(u_1)B(u_2) \\ C(u_1)C(u_2) & C(u_1)D(u_2) & D(u_1)C(u_2) & D(u_1)D(u_2) \end{pmatrix} \\
&= \begin{pmatrix} A(u_2)A(u_1) & B(u_2)A(u_1) & A(u_2)B(u_1) & B(u_2)B(u_1) \\ C(u_2)A(u_1) & D(u_2)A(u_1) & C(u_2)B(u_1) & D(u_2)B(u_1) \\ A(u_2)C(u_1) & B(u_2)C(u_1) & A(u_2)D(u_1) & B(u_2)D(u_1) \\ C(u_2)C(u_1) & D(u_2)C(u_1) & C(u_2)D(u_1) & D(u_2)D(u_1) \end{pmatrix} \begin{pmatrix} a_{12} & 0 & 0 & 0 \\ 0 & b_{12} & c_{12} & 0 \\ 0 & c_{12} & b_{12} & 0 \\ 0 & 0 & 0 & a_{12} \end{pmatrix} \\
\Rightarrow \quad &\Rightarrow \begin{pmatrix} a_{12}A_1A_2 & a_{12}A_1B_2 & a_{12}B_1A_2 & a_{12}B_1B_2 \\ b_{12}A_1C_2 + c_{12}C_1A_2 & b_{12}A_1D_2 + c_{12}C_1B_2 & b_{12}B_1C_2 + c_{12}D_1A_2 & b_{12}B_1D_2 + c_{12}D_1B_2 \\ c_{12}A_1C_2 + b_{12}C_1A_2 & b_{12}A_1D_2 + c_{12}C_1B_2 & c_{12}B_1C_2 + b_{12}D_1A_2 & c_{12}B_1D_2 + b_{12}D_1B_2 \\ a_{12}C_1C_2 & a_{12}C_1D_2 & a_{12}D_1C_2 & a_{12}D_1D_2 \end{pmatrix} \\
&= \begin{pmatrix} a_{12}A_2A_1 & b_{12}B_2A_1 + c_{12}A_2B_1 & c_{12}B_2A_1 + b_{12}A_2B_1 & a_{12}B_2B_1 \\ a_{12}C_2A_1 & b_{12}D_2A_1 + c_{12}C_2B_1 & c_{12}D_2A_1 + b_{12}C_2B_1 & a_{12}D_2B_1 \\ a_{12}A_2C_1 & b_{12}B_2C_1 + c_{12}A_2D_1 & c_{12}B_2C_1 + b_{12}A_2D_1 & a_{12}B_2D_1 \\ a_{12}C_2C_1 & b_{12}D_2C_1 + c_{12}C_2D_1 & c_{12}D_2C_1 + b_{12}C_2D_1 & a_{12}D_2D_1 \end{pmatrix}
\end{aligned}
$$

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

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.

quod erat dem■

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 $\boldsymbol{\mathcal{L}}$-matrix, defined by (70), its action on a tensor product of states can be written out as

$$\boldsymbol{\mathcal{L}}_{N,a} v \otimes \omega_+^i \begin{pmatrix} \lambda + \frac{i}{2} & \cdots \\ 0 & \lambda - \frac{i}{2} \end{pmatrix} v \otimes \omega_+^i, \ \forall v \in \mathbb{V}_a, \tag{91}$$

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.

$$\Omega^{\otimes N} = \bigotimes_{i=1}^{N} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \text{ where } \begin{aligned} A(u)\Omega^{\otimes N} &= \left(u + \tfrac{i}{2}\right)^N \Omega^{\otimes N}, \\ D(u)\Omega^{\otimes N} &= \left(u - \tfrac{i}{2}\right)^N \Omega^{\otimes N}, \end{aligned} \quad C(u)\Omega^{\otimes N} = 0. \tag{92}$$

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

$$|u_1, \cdots u_M\rangle = \prod_{j=1}^{M<N} B(u_j)\Omega^{\otimes N}, \text{ for some set } \{u_j\}_{j=1}^{M<N} \subset \mathbb{C}. \tag{93}$$

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 (93). 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,

$$\begin{aligned} A(v)\,|u_1, \cdots u_M\rangle &= \left[f(v-u_1)B(u_1)A(v) + g(v-u_1)B(v)A(u_1)\right] \prod_{j=2}^{M<N} B(u_j)\Omega^{\otimes N} \\ &= \left[\prod_{k=1}^{\ell} f(v-u_k)\right] \alpha^N(v)\,|u_1, \cdots u_M\rangle + \mathcal{O}(A \times D), \end{aligned} \tag{94}$$

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

(95)

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

Now, note that

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

$$\begin{aligned}\Rightarrow\sum_{k=1}^{\ell}\Big[\Big(g(v-u_k)\Big[\prod_{j\neq k}^{\ell}f(u_k-u_j)\Big]\alpha^N(u_k)\Big)+h(v-u_k)\Big(\prod_{j\neq k}^{\ell}h(u_k-u_j)\Big)\delta^N(u_j)\Big]&=0\\ \sum_{k=1}^{\ell}\Big[\Big(g(v-u_k)\Big[\prod_{j\neq k}^{\ell}f(u_k-u_j)\Big]\alpha^N(u_k)\Big)-g(v-u_k)\Big(\prod_{j\neq k}^{\ell}h(u_k-u_j)\Big)\delta^N(u_j)\Big]&=0\\ \sum_{k=1}^{\ell}g(v-u_k)\Big[\prod_{j\neq k}^{\ell}f(u_k-u_j)\Big]\alpha^N(u_k)+\prod_{j\neq k}^{\ell}h(u_k-u_j)\delta^N(u_j)\Big]&=0,\end{aligned}$$

which must be true for all k,

(96)
$$\Big[\prod_{j\neq k}^{\ell}f(v-u_k)\Big]\alpha^N(u_j)=\Big[\prod_{j\neq k}^{\ell}h(v-u_k)\Big]\delta^N(u_j).$$

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

(97)
$$\begin{aligned}\Big[\prod_{j\neq k}^{\ell}f(u_k-u_j)\Big]\alpha^N(u_j)&=\Big[\prod_{j\neq k}^{\ell}h(u_k-u_j)\Big]\delta^N(u_j)\\ \prod_{j\neq k}^{\ell}\frac{u_k-u_j-i}{u_k-u_j}\Big(u_j+\frac{i}{2}\Big)&=\prod_{j\neq k}^{\ell}\frac{u_k-u_j+i}{u_k-u_j}\Big(u_j-\frac{i}{2}\Big)\\ 0&=\prod_{j\neq k}^{\ell}\frac{u_k-u_j-i}{u_k-u_j}\Big(u_j+\frac{i}{2}\Big)-\prod_{j\neq k}^{\ell}\frac{u_k-u_j+i}{u_k-u_j}\Big(u_j-\frac{i}{2}\Big),\end{aligned}$$

which can be rearranged to yield the Bethe ansatz equations,

(98)
$$\boxed{\left(\frac{u_k+\frac{i}{2}}{u_k-\frac{i}{2}}\right)^N=\prod_{j\neq k}^{\ell}\frac{u_k-u_j+i}{u_k-u_j-i}},$$

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,

(99)
$$(A(v)+D(v))\,|u_1,\cdots u_M\rangle=\Big[\Big(\prod_{k=1}^{\ell}f(v-u_k)\Big)\alpha^N(v)+\Big(\prod_{k=1}^{\ell}h(v-u_k)\Big)\delta^N(v)\Big]\,|u_1,\cdots u_M\rangle$$
$$(A(v)+D(v))\,|u_1,\cdots u_M\rangle=\Big[\left(v+\frac{i}{2}\right)^N\prod_{k=1}^{\ell}\frac{v-u_k-i}{v-u_k}+\left(v-\frac{i}{2}\right)^N\prod_{k=1}^{\ell}\frac{v-u_k+i}{v-u_k}\Big]\,|u_1,\cdots u_M\rangle$$
$$\left(A\left(\frac{i}{2}\right)+D\left(\frac{i}{2}\right)\right)|u_1,\cdots u_M\rangle=i^N\prod_{k=1}^{\ell}\frac{u_k+\frac{i}{2}}{u_k-\frac{i}{2}}\,|u_1,\cdots u_M\rangle\equiv\lambda(\{u_k\}_{k=1}^{\ell<N})\,|u_1,\cdots u_M\rangle$$

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

(100)
$$P\,|u_1,\cdots u_\ell\rangle=\sum_{k=1}^{\ell}p(u_k)\,|u_1,\cdots u_\ell\rangle\ \text{ where }p(u_k)=\frac{1}{i}\log\frac{u_k+\frac{i}{2}}{u_k-\frac{i}{2}}.$$

Now, equation (89) 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,

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

Then, using equation (89), yields the eigenenergies

(101)
$$H\,|u_1,\cdots u_M\rangle=\left[\frac{i}{2}i\Big(\sum_{k=1}^{\ell}\frac{1}{u_k^2+\frac{1}{4}}-N\Big)-\frac{N}{2}\right]|u_1,\cdots u_M\rangle$$

or, equivalently,

(102)
$$\boxed{\mathbf{H}\,|u_1,\cdots u_M\rangle=\sum_{k=1}^{\ell}\epsilon(u_k)\,|u_1,\cdots u_M\rangle\ \text{ where }\epsilon(u_k)=-\frac{1}{2}\frac{1}{u_k^2+\frac{1}{4}}},$$

which concludes the Bethe solution for the XXX Heisenberg chain.

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

$$\mathbf{H} = J\sum_{j=1}^{L-1}(f_j^\dagger f_{j+1} + f_{j+1}^\dagger f_j) + \sum_{j=1}\lambda_j f_j^\dagger f_j, \quad \begin{matrix}\text{with the usual}\\ \text{conmutation rules}\end{matrix} \quad \begin{matrix}\{f_j, f_k\} = \{f_j^\dagger, f_k^\dagger\} = 0\\ \{f_j, f_k^\dagger\} = \delta_{jk}\end{matrix} \tag{103}$$

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[10]. Said Hamiltonian has open boundaries conditions since there is no hopping term across the boundary. Note that we can rewrite (103) as

$$\mathbf{H} = \sum_{i,j=1}^{L}\mathcal{M}_{ij}f_i^\dagger f_j \text{ with } \mathcal{M}_{ij} = \begin{cases}\lambda_i & \text{if } i=j\\ J & \text{if } j=i+1 \text{ or } i=j+1\\ 0 & \text{otherwise}\end{cases}, \quad \mathcal{M}\in \mathrm{GL}(L,\mathbb{R}), \tag{104}$$

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, (103) can be rewritten as

$$\mathbf{H} = \mathbf{f}^\dagger\mathcal{M}\mathbf{f}. \tag{105}$$

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 (103) to be rewritten as

$$\mathbf{H} = \mathbf{f}^\dagger A\mathcal{D}A^{\mathrm{T}}\mathbf{f} = \mathbf{d}^\dagger\mathcal{D}\mathbf{d} \tag{106}$$

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 (103)'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 (106) in terms of the lattices, as follows

$$\mathbf{H} = \sum_{k=1}^{L}\epsilon_k d_k^\dagger d_k, \tag{107}$$

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

$$\frac{d}{dt}d_k = i[\mathbf{H}, d_k], \tag{108}$$

and using that $d_k^2 = 0$, it turns out that (108)'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

[10] 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.

$$\mathcal{N}_{jk} = \langle d_j^\dagger d_k \rangle = \begin{cases} 0 \text{ or } 1 \text{ if } j = k \\ 0 \text{ if } j \neq k \end{cases}, \tag{109}$$

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,

$$\mathcal{N}_{jk} = \langle d_j^\dagger d_k \rangle_{\text{th}} = \frac{1}{1 + e^{\beta\epsilon_k + \mu}} \delta_{jk}. \tag{110}$$

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

$$\langle f_j^\dagger f_k \rangle = \sum_{k,l=1}^{n} A_{jk} A_{jl} \langle d_k^\dagger d_l \rangle. \tag{111}$$

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

$$\langle f_j^\dagger(t) f_k(t) \rangle = \sum_{k,l=1}^{L} e^{i(\epsilon_k - \epsilon_l)t} A_{jk} A_{jl} \langle d_k^\dagger d_l \rangle, \tag{112}$$

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^0_{ion} be the particle density of the ion jellium and let $\rho^0_{\text{el}} = Z\rho^0_{\text{ion}}$ 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

$$\rho_{\text{ion}}(x^\mu) = \rho^0_{\text{ion}} + \delta\rho_{\text{ion}}(\mathbf{x})e^{-i\Omega t}. \tag{113}$$

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.

$$\nabla\cdot\mathbf{E} = \frac{Ze}{\epsilon_0}\delta\rho_{\text{ion}} \qquad\qquad \Rightarrow \nabla\cdot\mathfrak{f} = \frac{Z^2e^2\rho^0_{\text{ion}}}{\epsilon_0}\delta\rho_{\text{ion}}, \tag{114}$$

where $\mathfrak{f} = Ze\rho_{\text{ion}}\mathbf{E} \approx Ze\rho^0_{\text{ion}}\mathbf{E}$ 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

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

since the velocity is already a small quantity. Differentiating this with respect to time once and again and using Newton's second law $\mathfrak{t} = M\rho_{\text{ion}}\partial_t\mathbf{v}$ 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

$$\begin{aligned} \langle \mathbf{A} \rangle = \frac{1}{\mathcal{Z}_0} \mathrm{Tr}_{\,\mathbb{H}} \rho_0 \mathbf{A} = \frac{1}{\mathcal{Z}_0} \sum_{n \in \Lambda} \langle n | \, \mathbf{A} \, | n \rangle \, e^{-\beta E_n} \quad &\text{where } \{|n\rangle\}_{n\in\Lambda} \text{ is a complete} \\ \rho_0 = e^{-\beta \mathbf{H}_0} = \sum_{n \in \Lambda} |n\rangle \, \langle n| \, e^{-\beta E_n}, \qquad &\text{set of eigenstates.} \end{aligned} \tag{115}$$

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

$$\mathbf{H}(t) = \mathbf{H}_0 + \mathbf{H}'(t)\theta(t - t_0) \tag{116}$$

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

$$\langle \mathbf{A}(t) \rangle = \frac{1}{\mathcal{Z}_0} \mathrm{Tr}_{\,\mathbb{H}} \rho(t) \mathbf{A} = \frac{1}{\mathcal{Z}_0} \sum_{n \in \Lambda} \langle n(t)| \, \mathbf{A} \, |n(t)\rangle \, e^{-\beta E_n} \tag{117}$$

$$\rho_0 = e^{-\beta \mathbf{H}_0} = \sum_{n \in \Lambda} |n(t)\rangle \, \langle n(t)| \, e^{-\beta E_n}, \tag{118}$$

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

$$|n(t)\rangle = e^{-i\mathbf{H}_0 t} \, |n(t)\rangle_I = e^{-i\mathbf{H}_0 t} \mathcal{U}(t, t_0) \, |\hat{n}(t_0)\rangle \, , \tag{119}$$

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

(120)

$$\begin{aligned}\langle \mathbf{A}(t)\rangle &= \frac{1}{\mathcal{Z}_0}\sum_{n\in\Lambda}\langle n(t)|\,\mathbf{A}\,|n(t)\rangle\, e^{-\beta E_n}\\ &= \frac{1}{\mathcal{Z}_0}\sum_{n\in\Lambda}\langle n(t_0)|\left(\mathbb{1}-i\int_{t_0}^{t}dt'\mathbf{H}'(t')\right)e^{-i\mathbf{H}_0t_0}\mathbf{A}e^{i\mathbf{H}_0t_0}\left(\mathbb{1}-i\int_{t_0}^{t}dt'\mathbf{H}'(t')\right)|n(t_0)\rangle\, e^{-\beta E_n}+\mathcal{O}(\mathbf{H}'^2)\\ &= \frac{1}{\mathcal{Z}_0}\sum_{n\in\Lambda}\langle n(t_0)|\,\mathbf{A}\,|n(t_0)\rangle - \frac{i}{\mathcal{Z}_0}\int_{t_0}^{t}dt'\sum_{n\in\Lambda}e^{-\beta E_n}\left(e^{-i\mathbf{H}_0t_0}\mathbf{A}e^{i\mathbf{H}_0t_0}\mathbf{H}'(t')-\mathbf{H}'(t')e^{-i\mathbf{H}_0t_0}\mathbf{A}e^{i\mathbf{H}_0t_0}\,|n(t_0)\rangle\right)\\ &= \langle\mathbf{A}\rangle_0+\int_{t_0}^{t}\frac{dt'}{i}\langle[\mathbf{A}(t),\mathbf{H}'(t')]\rangle_0,\end{aligned}$$

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.

(121)
$$\boxed{\delta\langle\mathbf{A}(t)\rangle=\int_{t_0}^{\infty}dt'\mathcal{C}^R_{AH'}(t,t')e^{-\eta(t-t')}\text{ where }\mathcal{C}^R_{AH'}(t,t')=-i\theta(t-t')\langle[\mathbf{A}(t),\mathbf{H}'(t')]\rangle_0},$$

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[11].

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[12]. Therefore, consider the perturbation Hamiltonian $\mathbf{H}'$, which can be rewritten in terms of its Fourier components

(123)
$$\mathbf{H}'(t)=\int_{\Omega\subset\mathbb{R}}\frac{d\omega}{2\pi}e^{-i\omega t}\mathbf{H}'_\omega,$$

such that the retarded correlation function becomes

(124)
$$\mathcal{C}^R_{AH'}(t,t')=\int_{\mathbb{R}}\frac{d\omega}{2\pi}e^{-i\omega t'}\mathcal{C}^R_{AH'_\omega}(t-t')\quad\begin{array}{l}\text{since }\langle[\mathbf{A}(t),\mathbf{H}'(t')_{\omega'}]\rangle_0\text{ only depends}\\ \text{on the difference between }t\text{ and }t'.\end{array}$$

Therefore, inserting this result into the Kubo formula yields

[11] 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.

[12] Consider the L^1-space, ie. the space of all integrable functions on the real line. Then the Fourier transform and the Fourier-anti trasnform can be defined as

(122)
$$\mathcal{F}[f(t)](\omega)=f(\omega)=\int_{\mathbb{R}}dte^{i\omega t}f(t)\text{ and }\mathcal{F}^{-1}[f(\omega)](t)=f(t)=\int_{\mathbb{R}}\frac{d\omega}{2\pi}e^{-i\omega t}f(\omega)$$

$$
\begin{aligned}
\delta\langle \mathbf{A}(t)\rangle &= \int_{t_0}^{\infty} dt' \mathcal{C}^R_{AH'_\omega}(t,t') e^{-\eta(t-t')} \\
&= \int_{\mathbb{R}} dt' \int_{\mathbb{R}} \frac{d\omega}{2\pi} e^{-i\omega t} e^{-i(\omega+i\eta)(t'-t)} \mathcal{C}^R_{AH'_\omega}(t-t') \\
&= \int_{\mathbb{R}} \frac{d\omega}{2\pi} e^{-?`i\omega t} \Big(\int_{\mathbb{R}} d(t'-t) e^{-i(\omega+i\eta)(t'-t)} \mathcal{C}^R_{AH'_\omega}(t-t') \Big) \\
&= \int_{\mathbb{R}} \frac{d\omega}{2\pi} e^{-i\omega t} \mathcal{C}^R_{AH'_\omega}(\omega;\eta)
\end{aligned} \tag{125}
$$

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

$$
\begin{aligned}
\int_{\mathbb{R}} dt e^{i\nu t} \delta\langle \mathbf{A}(t)\rangle &= \int_{\mathbb{R}} dt e^{i\nu t} \int_{\mathbb{R}} \frac{d\omega}{2\pi} e^{-i\omega t} \mathcal{C}^R_{AH'_\omega}(\omega;\eta) \\
\langle \mathbf{A}_\nu\rangle &= \int_{\mathbb{R}} \frac{d\omega}{2\pi} \int_{\mathbb{R}} dt e^{i(\nu-\omega)t} \mathcal{C}^R_{AH'_\omega}(\omega;\eta) \\
&= \int_{\mathbb{R}} d\omega \delta(\nu-\omega) \mathcal{C}^R_{AH'_\omega}(\omega;\eta) \\
&= \mathcal{C}^R_{AH'_\nu}(\nu;\eta)
\end{aligned} \tag{126}
$$

$$
\Rightarrow \delta\langle \mathbf{A}_\omega\rangle = \mathcal{C}^R_{AH'_\omega}(\omega) \text{ with } \mathcal{C}^R_{AH'_\omega}(\omega) = \int_{\mathbb{R}} dt e^{i\omega t} e^{-\eta t} \mathcal{C}^R_{AH'_\omega}(t), \tag{127}
$$

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.[13]

$$
J^\alpha_e(x^\mu) = \int_{\Omega\subset\mathbb{R}^4} dx^\mu \sum_{\alpha\beta} \sigma_\beta(x^\mu, y^\mu) E^\beta(y^\mu), \tag{128}
$$

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,

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

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

[13] 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 $g_{\mu\nu} = \delta_{\mu\nu}$

$$\mathbf{J} = \mathbf{J}^{\nabla}(\mathbf{x}) + \frac{e}{m}\mathbf{A}(\mathbf{x})\rho(\mathbf{x}). \tag{129}$$

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

$$\partial_t \xrightarrow{F} -i\omega \quad \text{then} \;\; \mathbf{A}_{\text{ext}}(\mathbf{x},\omega) \xrightarrow{F} \tfrac{1}{i\omega}\mathbf{E}_{\text{ext}}(\mathbf{x},\omega) \;\; \Rightarrow \mathbf{H}_{\text{ext},\omega} = \frac{e}{i\omega}\int_{\Omega\subset\mathbb{R}^3} d\mathbf{x}\, \mathbf{J}(\mathbf{x})\cdot\mathbf{E}_{\text{ext}}(\mathbf{x},\omega). \tag{130}$$

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 onyly 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

$$J_e^\alpha(\mathbf{x},\omega) = \int_{\Omega\subset\mathbb{R}^3} d\mathbf{y} \;\sum_\beta \sigma_{\alpha\beta}(\mathbf{x},\mathbf{y},\omega)E^\beta(\mathbf{y},\omega). \tag{131}$$

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

4. .