Schwinger boson mean-field theory to the Heisenberg model

Yi-Ming Ding[1]

[1]*Department of Physics, School of Science, Westlake University*

(Dated: June 30, 2024)

The Heisenberg model serves as a foundational paradigm for studying magnetism in condensed matter physics, while exact solutions are not generally accessible for those systems with mixed spins, frustration or in higher dimensions. This term paper reviews one of popular methodologies called the Schwinger boson mean-field theory (SBMFT) to address the problem above, which treats the Heisenberg model within the Schwinger boson representation and at the mean-field level.

I. INTRODUCTION

Mean-field approximation (MFA) may be the most frequently used method in condensed matter theory to investigate many-body interacting systems. By substituting some quantity with the corresponding expectation value, we are able to decouple many-body interactions into fewer-body interactions. As MFA ignores the effect of fluctuations, one can expect that it could result in some discrepancy when the fluctuations are nonnegligible. Nevertheless, it would be acceptable if an MFA gives excellent and qualitative explanations, not necessarily quantitive. In this situation, it can extract and grasp the fine and core structure of system.

Another technique often in conjunction with MFA is to introduce some **auxiliary particles**. Mathematically, this converts the representation of the model, which may also help decoupling many-body interactions and possibly enables a better MFA to solve the problem. Common formalisms of auxiliary particle include the slave boson/fermion representation, and the **Schwinger boson (SB) representation**. Therefore, just as the name implies, the Schwinger boson mean-field theory (SBMFT) is basically a mean-field approach within the SB representation, and we will introduce how to apply it to the **Heisenberg model** in this paper.

The natural question following is why do we use the SB representation or what are the merits of it? Part of the reasons come from its comparison with the **Holstein–Primakoff (HP) representation**, which is used in the conventional spin-wave theory, and we briefly review it here. Recall that $\mathrm{SU}(2)$ describes the internal symmetry of the Heisenberg model, whose algebra is $[\hat{S}_i^\alpha, \hat{S}_j^\beta] = \delta_{ij}(i\varepsilon_{\alpha\beta\gamma}\hat{S}_i^\gamma)$, where the subscript i and j label two sites on a lattice. Then the HP transformation rewrites the spin operators by

$$\hat{S}_i^z = S - \hat{n}_i, \quad \hat{S}_i^+ = \sqrt{2S - \hat{n}_i}\hat{h}_i, \quad \hat{S}_i^- = h_i^\dagger\sqrt{2S - \hat{n}_i} \tag{1}$$

where $n_j = \hat{h}_i^\dagger \hat{h}_i$ denotes the number operator and $\hat{h}_i^\dagger$ ($\hat{h}_i$) denotes the bosonic creation (annihilation) operator, satisfying the bosonic commutative relations $[\hat{h}_i, \hat{h}_j^\dagger] = \delta_{ij}$ and $[\hat{h}_i, \hat{h}_j] = [\hat{h}_i^\dagger, \hat{h}_j^\dagger] = 0$ for site i and site j on the lattice. In addition, a constraint

$$0 \leq \hat{n}_i \leq 2S \tag{2}$$

should be introduced, which imposes the boundaries in the Hilbert space. Two drawbacks can be directly noticed from the first glance of the HP representation. One is the sqaure root involved in the transformation (1), and another is the non-holonomic constraint (2), both of which are inconvenient to handle in practical calculations. Besides, the HP representation or the spin-wave theory relies on a magnetized ground state, which also limits its application. As we will see later, none of the problems mentioned above occurs in the SB representation. It can be applied to both ordered and disordered phases and general mixed spins systems. Unlike the large-S limit taken in spin-wave theory for the $1/S$ expansion, the SB representation preserves the rotation symmetry of the model.

One more thing we would like to discuss in this paper is the **large-N approximation**. This is an interesting and magical method used in both particle physics and condensed matter physics. Taking the Heisenberg model as an example, first we generalize the internal symmetry from $\mathrm{SU}(2)$ to $\mathrm{SU}(N)$, which turns out to be easier to solve. Then after deriving the results in the limit of large N we directly obtain the result of $\mathrm{SU}(2)$ by crudely substituting N with 2. The results can be qualitatively true or completely wrong, depending on the model. Roughly thinking, the large-N approximation shares some similarity with the large-S approximation as the total spin at each latice site is NS ($N \to \infty$), but this is not a simple classical limit. As we will see later, it is easy to generalize the SB representation to $\mathrm{SU}(N)$, therefore we can investigate the so-called $\mathrm{SU}(N)$ Heisenberg model by SBMFT as well.

This paper is organized as follows. We introduce the SB representation of $\mathrm{SU}(2)$ and the generalized $\mathrm{SU}(N)$ in Sec. II, for both the case of **ferromagnetic/ferromagnet (FM)** and **antiferromagnetic/antiferromaget (AFM)**. In Sec. III, I will illustrate SBMFT with the example of a d-dimensional $\mathrm{SU}(2)$ FM Heisenberg model, which is though a little bit trivial, exhibiting a standard procedure of SBMFT. Then in Sec.IV, we move to a more interesting example, the **ferrimagnetic (FiM)** chain constituted of two different spins. At the last, a sketch of the main results of the $\mathrm{SU}(N)$ AFM model will be introduced in Sec.V, and in Sec. VI, a summary. All references are listed in the last page.

II. SCHWINGER BOSON REPRESENTATION

A. SU(2)

We start with SU(2), and the SB transformation is given by

$$\hat{S}_i^z = \frac{1}{2}(\hat{a}_i^\dagger \hat{a}_i - \hat{b}_i^\dagger \hat{b}_i), \quad \hat{S}_i^+ = \hat{a}_i^\dagger \hat{b}_i, \quad \hat{S}_i^- = \hat{b}_i^\dagger \hat{a}_i, \tag{3}$$

where two flavors of bosons (**Schwinger bosons**), are introduced, satisfying

$$\begin{aligned}
&[\hat{a}_i, \hat{a}_j^\dagger] = [\hat{b}_i, \hat{b}_j^\dagger] = \delta_{ij}, \\
&[\hat{a}_i, \hat{a}_j] = [\hat{a}_i^\dagger, \hat{a}_j^\dagger] = [\hat{b}_i, \hat{b}_j] = [\hat{b}_i^\dagger, \hat{b}_j^\dagger] = 0 \\
&[\hat{a}_i, \hat{b}_j] = [\hat{a}_i^\dagger, \hat{b}_j] = [\hat{b}_i, \hat{a}_j] = [\hat{b}_i^\dagger, \hat{a}_j] = 0.
\end{aligned} \tag{4}$$

For each site, a holonomic constraint $\hat{a}_i^\dagger \hat{a}_i + \hat{b}_i^\dagger b_i = 2S$ is imposed to ensure the total spin, which enables us to include it using some Lagrange multiplier. Mathematically, SB representation can give all irreducible representation of SU(2), and each of them is uniquely labeled by the total number of Schwinger bosons (shown in Fig. 1).

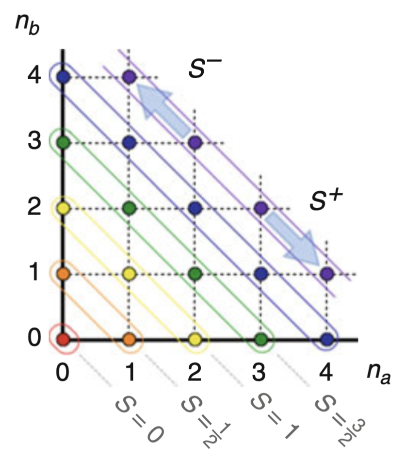

FIG. 1. A schematic diagram of the Schwinger boson representation, where n_a and n_b correspond to the particle numbers of $\hat{a}$ and $\hat{b}$ bosons, respectively. Figure from Ref. [6].

Within the SB representation, we can actually write the Heisenberg model in a very compact form. We first consider the FM case, and the corresponding Hamiltonian is

$$\hat{H}_{\text{FM}} = -J \sum_{\langle ij \rangle} \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j = -J \sum_{\langle ij \rangle} \left[\hat{S}_i^z \hat{S}_j^z + \frac{1}{2}(\hat{S}_i^+ \hat{S}_j^- + \hat{S}_i^- \hat{S}_j^+) \right], \tag{5}$$

where $J > 0$ and $\langle ij \rangle$ denotes the nearest neighbor site-pair. We define the **FM bond operator**

$$\hat{F}_{ij} := \hat{a}_i^\dagger \hat{a}_j + \hat{b}_i^\dagger \hat{b}_j \tag{6}$$

then

$$\hat{H}_{\mathrm{FM}} = -\frac{J}{2}\sum_{\langle ij\rangle}(:\hat{F}_{ij}^{\dagger}\hat{F}_{ij}: -2S^2) \tag{7}$$

where :: denotes the normal ordering. Notice that in Eq. (7), the interaction between two adjacent sites is completely described by $\hat{F}_{ij}$, and this is the reason why we call $\hat{F}_{ij}$ the bond operator in this context.

Similarly, for the SU(2) antiferromagnetic (AFM) Heisenberg model

$$\hat{H}_{\mathrm{AFM}} = J\sum_{\langle ij\rangle}\hat{\mathbf{S}}_i\cdot\hat{\mathbf{S}}_j = J\sum_{\langle ij\rangle}\left[\hat{S}_i^z\hat{S}_j^z + \frac{1}{2}(\hat{S}_i^+\hat{S}_j^- + \hat{S}_i^-\hat{S}_j^+)\right], \tag{8}$$

define the **AFM bond operator**

$$\hat{A}_{ij} := \hat{a}_i\hat{b}_j - \hat{b}_i\hat{a}_j \tag{9}$$

then

$$\hat{H}_{\mathrm{AFM}} = -\frac{J}{2}\sum_{\langle i\rangle j}(\hat{A}_{ij}^{\dagger}\hat{A}_{ij} - 2S^2) \tag{10}$$

Comparing the two bond operators defined in Eq. (6) and Eq. (9), we find that the two flavors of bosons couple together in the AFM case. On another hand, if we generalize the SB representation to SU(N), one should expect that N flavors of bosons would be introduced at each site. Therefore, the coupling of different bosons in the AFM case is less tractable than that of the FM case. For a bipartite lattice, this can typlically be fixed by some gauge transformation or rotation of the coordinate on one of the sublattice. Specifically, assume the lattice includes two sublattice A and B, then taking

$$\hat{a}_j \to -\hat{b}_j, \quad \hat{b}_j \to \hat{a}_j \tag{11}$$

where $i\in A$ and $j\in B$, we can rewrite

$$\hat{A}_{ij} \to \hat{a}_i\hat{a}_j + \hat{b}_i\hat{b}_j \tag{12}$$

Apparently, the functional form of the Hamiltonian (10) is unchanged under the transformation (11). In this paper, we restrict our discussions on bipartite lattice. For the frustrated models, one may refer to [N. Read and S. Sachdev, Phys. Rev. Lett. 66, 1773], where the Schwinger boson representation is extended to the Sp(N) group.

B. SU(N)

The algebra of SU(N) is

$$[\hat{S}_i^{\mu\nu}, \hat{S}_j^{\mu',\nu'}] = \delta_{ij}(\hat{S}_i^{\mu\nu'}\delta^{\mu'\nu} - \hat{S}_i^{\mu'\nu}\delta^{\mu\nu'}) \tag{13}$$

with which we define the **generalized SU(N) Heisenberg model**

$$\hat{H}_{\mathrm{N}} = \sigma \frac{J}{N} \sum_{\langle ij \rangle} \left(\sum_{\mu\nu} \hat{S}_i^{\mu\nu} \hat{S}_j^{\nu\mu} - NS^2 \right) \tag{14}$$

where $\sigma = \pm 1$ is for denoting the FM and AFM interaction, and $\mu,\, \nu \in 1, \ldots N$.

The generalization (14) may not look straightforward. However, by identifying

$$\hat{S}_i^{12} \equiv \hat{S}_i^{+}, \quad \hat{S}_i^{21} \equiv \hat{S}_i^{-}, \quad \hat{S}_i^{11} \equiv \hat{S}_i^{z} + S, \quad \hat{S}_i^{22} \equiv \hat{S}_i^{z} - S \tag{15}$$

it is easy to verify that Eq. (14) can reduce to the standard SU(2) Heisenberg models (5) and (8).

Akin to SU(2), by introducing N different flavors of bosons $\hat{b}_\mu$ with $\mu = 1, \ldots, N$ at each site, we can define the SB representation of SU(N), which is

$$\hat{S}_i^{\mu\nu} = \hat{b}_{i\mu}^\dagger \hat{b}_{i\nu} \tag{16}$$

If the interaction is AFM, we keep the transformation (16) on sublattice A and further define $\hat{S}_j^{\mu\nu} = -\hat{b}_{j\mu}^\dagger \hat{b}_{j\nu}$ on sublattice B.

The holonomic constrant now imposed at each site is

$$\sum_{\mu=1}^{N} \hat{b}_{i\mu}^\dagger \hat{b}_{i\mu} = NS := \hat{n}_s \tag{17}$$

where the total number of Schwinger bosons n_s specifies different irreducible representations of SU(N).

For $\sigma = -1$, i.e. the FM case, in Eq. (14), define the SU(N) FM bond operator

$$\hat{F}_{ij} := \sum_{\mu=1}^{N} \hat{b}_{i\mu}^\dagger \hat{b}_{i\mu} \tag{18}$$

which results in the SU(N) FM Heisenberg model

$$\hat{H}_{\text{FM-N}} = -\frac{J}{N} \sum_{\langle ij \rangle} (:\hat{F}_{ij}^\dagger \hat{F}_{ij}: - NS^2) \tag{19}$$

Similarly, for $\sigma = 1$, i.e. the AFM case, in Eq. (14), define the SU(N) AFM bond operator

$$\hat{A}_{ij} := \sum_{\mu=1}^{N} \hat{b}_{i\mu} \hat{b}_{j\mu} \tag{20}$$

and the compact form of the SU(N) AFM Heisenberg model is given by

$$\hat{H}_{\text{AFM-N}} = -\frac{J}{N} \sum_{\langle ij \rangle} (\hat{A}_{ij}^\dagger \hat{A}_{ij} - NS^2) \tag{21}$$

III. EXAMPLE 1: THE SU(2) FM HEISENBERG MODEL

In the following sections, we will introduce three examples of SBMFT with different types of Heisenberg models.

As we have introduced in Sec. I, the ultimate purpose of the large-N expansion is to investigate the physical case of small N. In the initial work of A. Auerbach and D. P. Arovas (see Ref. [1, 2]), they discuss both the SB representation and the generalized SU(N) Heisenberg model. They adopted a functional integral approach in conjunction with the steepest descent approximation to derive their results. Their method, at least to me, seems complicated and requires a lot of mathematical techniques. Opportunely, it soon turned out that a direct MFA is sufficient to our concerned SU(2) Heisenberg model, offering qualiatively good explanations for many properties. Therefore, we prefer a direct MFA rather than Auerbach and Arovas's old way in this paper. In addition, as a review paper, we will also give a quick sketch of the functional integral approach in the last example without details.

A. Mean-field Hamiltonian

Recall the SU(2) FM Heisenberg model

$$\hat{H}_{\text{FM}} = -2J \sum_{\langle ij \rangle} : \hat{F}_{ij}^{\dagger} \hat{F}_{ij} : + \frac{VzJ}{2} S^2 \tag{22}$$

where z is the coordination number, and V is the number of sites (volume). Notice that we have defined the bond operator as

$$\hat{F}_{ij} = \frac{1}{2} \sum_{\mu} \hat{b}_{i\mu}^{\dagger} \hat{b}_{j\mu}, \quad \mu = \uparrow, \downarrow \tag{23}$$

where the arrows $\{\uparrow, \downarrow\}$ are used to label two flavors of bosons rather than the real spins.

The translation and rotation symmetry of the lattice enable us to use a single λ Lagarange multiplier for all the sites, which play the role of chemical potential, in the FM case to include the constriant $\sum_{\mu} \hat{b}_{i\mu}^{\dagger} \hat{b}_{i\mu} = 2S$ at each site. This is not true for the AFM case, especially in a frustrated system, and we may need several λ.

After a Hatree-Fock decomposition, we can obtain the mean-field Hamiltonian

$$\hat{H}_{\text{FM-HF}} = \lambda \sum_{i} \left[\sum_{\mu} \hat{b}_{i\mu} \hat{b}_{i\mu} - 2S \right] - 2J \sum_{\langle ij \rangle} \left[\langle \hat{F}_{ij}^{\dagger} \rangle \hat{F}_{ij} + \langle \hat{F}_{ij} \rangle \hat{F}_{ij}^{\dagger} \right] + 2J \sum_{\langle ij \rangle} \langle \hat{F}_{ij}^{\dagger} \rangle \langle \hat{F}_{ij} \rangle + \frac{VzJ}{2} S^2 \tag{24}$$

Notice that the two channels of $\uparrow$ and $\downarrow$ are automatically decoupled.

At the mean-field level, we further assume $\langle F_{ij}^{\dagger} \rangle = \langle \hat{F}_{ij} \rangle = \tilde{F}$, which is a real and uniform number for all sites. Remember that $\hat{F}_{ij}$ accounts for the interaction between site i and site j, therefore a non-zero value of $\tilde{F}$ signfies a short-range FM correlation.

Then we apply a Fourier transform to achieve the mean-field Hamiltonian in the momentum space, which is

$$\hat{H}_{\text{FM-MF}} = VJS^2 - 2\lambda SV + 2VJ\tilde{F}^2 + \sum_{\mathbf{k}\mu} \omega_{\mathbf{k}} \hat{b}^{\dagger}_{\mathbf{k}\mu} \hat{b}_{\mathbf{k}\mu} \tag{25}$$

where

$$\omega_{\mathbf{k}} = J\tilde{F}(\varepsilon_{\mathbf{k}} + \Lambda) = J\tilde{F}z(\varepsilon_{\mathbf{k}} - 1) + \lambda \tag{26}$$

gives the **dispersion relation** of the Schwinger bosons, with

$$\Lambda = \lambda / J\tilde{F}z - 1, \quad \varepsilon_{\mathbf{k}} = \frac{1}{z}\sum_{\delta}(1 - e^{i\mathbf{k}\cdot\delta}) \tag{27}$$

The latter one sums over all lattice vector basis δ.

With some algebras, we can obtain the free energy per site, denoted by f, satisfying

$$f = -\frac{Jz}{2}S^2 - 2Jz\lambda S\tilde{F} - Jz(\tilde{F} - S)^2 - \frac{2}{N\beta}\sum_{\mathbf{k}} \ln(1 + n_{\mathbf{k}}) \tag{28}$$

where $n_{\mathbf{k}} = (e^{\omega_{\mathbf{k}}/k_B T} - 1)^{-1}$ is the Bose occupation factor.

The relations among the parameters involved in our calculations including $\tilde{F}$ can be obtained by minimizing f, resulting in

$$S = \frac{1}{V}\sum_{\mathbf{k}} n_{\mathbf{k}} \tag{29}$$

$$\tilde{F} = S - \frac{1}{V}\sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} n_{\mathbf{k}} \tag{30}$$

which are called the **self-consistent equations**.

By replacing the summations in the Eq. (29) and (30) with integrals in the first Brillouin zone, we have

$$S = \int \frac{d^d k}{(2\pi)^d} \frac{1}{e^{\tilde{\beta}(\varepsilon_{\mathbf{k}} + \Lambda)} - 1} \tag{31}$$

$$\tilde{F} = S - \int \frac{d^d k}{(2\pi)^d} \frac{\varepsilon_{\mathbf{k}}}{e^{\tilde{\beta}(\varepsilon_{\mathbf{k}} + \Lambda)} - 1} \tag{32}$$

where d is the spatial dimension and

$$\tilde{\beta} := J\tilde{F}z/(k_B T). \tag{33}$$

is some effective inverse temperature.

B. Criticality for $d > 2$

For the zero temperature $T = 0$, Eq. (31) has no solution for any dimensionality.

If $d = 1,2$, Eq. (31) has a solution for any $T > 0$. Since the system is disordered in infinite high temperature (the thermal fluctuations can destroy any order), the continuity of the solution suggests that the system keeps disordered as long as $T > 0$. This result is consistent with the **Mermin-Wagner theorem** as the symmetry is unbroken.

Interesting thing happens when $d > 2$. There exists a temperature T_c, below which there is no solution, and above which there is a solution. We expect T_c should to be the critical point, but we need more clues to affirm this. If T_c is indeed the critical point, then can we explain an ordered-disordered phase transition with it? Furthermore, as the region $T < T_c$ has no solution, can we extend our theory to that region?

In fact, accompanied with T_c, the quantity Λ defined in Eq. (27) can actually remove our doubts. For $T \geq T_c$, $\Lambda \geq 0$, and $\Lambda \to 0^+$ as $T \to T_c^+$. Recall that $\omega_{\mathbf{k}} = J\tilde{F}(\varepsilon_{\mathbf{k}} + \Lambda)$ defined in Eq. (26), then the gap is finite when $T > T_c$, and closes to zero as $T \to T_c^+$, indicating $T = T_c$ is to a gapless point. This observation means if T_c is a critical point, Λ can reflect the critical behaviors. Nevertheless, we assume T_c is a critical point first.

For $T = T_c$ or $\tilde{\beta} = \tilde{\beta}_c$, since $\Lambda = 0$, Eq. (31) and (32) become

$$S = \int \frac{d^d k}{(2\pi)^d} \frac{1}{e^{\tilde{\beta}_c \varepsilon_{\mathbf{k}}} - 1} \tag{34}$$

$$\tilde{F}_c = S - \int \frac{d^d k}{(2\pi)^d} \frac{\varepsilon_{\mathbf{k}}}{e^{\tilde{\beta}_c \varepsilon_{\mathbf{k}}} - 1} \tag{35}$$

For convenience, we can define an analytical function

$$I(\tilde{\beta}) := \int \frac{d^d k}{(2\pi)^d} \frac{1}{e^{\tilde{\beta} \varepsilon_{\mathbf{k}}} - 1} \tag{36}$$

which reduces to Eq. (34) when $\tilde{\beta} = \tilde{\beta}_c$.

To extract the term that reflects the criticality, we consider the substraction

$$I(\tilde{\beta}) - I(\tilde{\beta}_c) \tag{37}$$

for $\tilde{\beta} < \tilde{\beta}_c$ ($T > T_c$) and $\tilde{\beta} \to \tilde{\beta}_c^-$ ($T \to T_c^+$).

With some algebra and expanding $I(\tilde{\beta}) - I(\tilde{\beta}_c)$ by small energy, we can achieve

$$\Lambda \sim (\tilde{\beta} - \tilde{\beta}_c)^s \tag{38}$$

i.e. Λ decay as a power law at the critical point for some s. Specifically,

$$\Lambda \sim \begin{cases} (\tilde{\beta} - \tilde{\beta}_c)^{\frac{2}{(d-2)}} & 2 < d < 4 \\ -\dfrac{\tilde{\beta}_c - \tilde{\beta}}{\ln(\tilde{\beta}_c - \tilde{\beta})}, & d = 4 \\ 1, & d > 4 \end{cases} \tag{39}$$

One may notice that so far we have only used Eq. (31) without discussing Eq. (32) regarding $\tilde{F}$. Actually, different from $\Lambda_c = 0$ (the value of Λ at T_c), $\tilde{F}_c$ is finite, and $\tilde{F} - \tilde{F}_c \sim t$, where $t := (T - T_c)/T_c$. With some calculations, this can also gives the same behavior of Λ as above.

To further explore the criticality, we consider the spin-spin correlation, which in our SBMFT, can be written as

$$G(\mathbf{r}) := \langle \hat{\mathbf{S}}(\mathbf{0}) \cdot \hat{\mathbf{S}}(\mathbf{r}) \rangle = \frac{3}{2}\left[S\delta_{\mathbf{r0}} + |g(\mathbf{r})|^2 \right] \tag{40}$$

where

$$g(\mathbf{r}) := \int \frac{d^d k}{(2\pi)^d} \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{e^{\tilde{\beta}(\varepsilon_{\mathbf{k}}+\Lambda)} - 1} \tag{41}$$

Before we move on, let us pause a little bit here. Notice that

$$G(\mathbf{0}) = \frac{3}{2} S(S+1) \neq S(S+1) \tag{42}$$

This discrepancy actually comes from the overcounting in the MFA particle numbers on each site, i.e. $\langle \hat{F}_{ii} \rangle \sim \tilde{F}$. Intutively thinking, this relaxtion of praticle number constraint make two flavors of bosons independent, resulting the overcounting.

The correlation length can then be obatined from the correlation function, yielding

$$\xi \sim \begin{cases} t^{-\nu} = t^{-\frac{1}{d-2}} & 2 < d < 4 \\ \text{same as the spherical model,} & d \geq 4 \end{cases} \tag{43}$$

Similarly, for the susceptibility, one can obtain

$$\chi \sim t^{-\gamma}, \quad \gamma = \frac{4-d}{d-2}, \quad 2 < d < 4 \tag{44}$$

In particular at $d = 3$, $\nu = 1$ and $\gamma = 1$.

All the behaviors above are *same to an idea Bose gas* when approaching the critical point of the Bose-Einstein condenstation (BEC), but here is an interacting Bose gas, and the interaction comes from $\tilde{F}$. Therefore, after we change the representation from spin to boson, the ordered phase of spins become the condensation of Schwinger bosons. Thank Einstein, we can now convince ourselves.

C. Bose-Einstein condensation and symmetry breaking

We can assume BEC occurs equivalently for the channel $\uparrow$ and channel $\downarrow$ at T_c, i.e.

$$\langle \hat{b}_{i\uparrow} \hat{b}_{i\uparrow} \rangle = \langle \hat{b}_{i\downarrow} \hat{b}_{i\downarrow} \rangle = S, \quad \langle \hat{S}_i^z \rangle = 0 \tag{45}$$

according to the SB transformation (3). Then our target is to find a solution satisfying the conditions above.

Before turning Eq. (31) into an integral, we can extract the the mode of $\mathbf{k}=\mathbf{0}$ for $\omega_{\mathbf{k}}$, then

$$S=\rho+\int\frac{d^dk}{(2\pi)^d}\frac{1}{e^{\tilde{\beta}(\varepsilon_{\mathbf{k}}+\Lambda)}-1} \tag{46}$$

where we have defined

$$\rho:=\frac{n_{\mathbf{0}}}{V}=\frac{\langle\hat{b}^{\dagger}_{\mathbf{0}\uparrow}\hat{b}_{\mathbf{0}\uparrow}\rangle}{V}=\frac{1}{V(e^{\tilde{\beta}\Lambda}-1)} \tag{47}$$

If $T>T_c$, $\rho=0$ in the thermodynamic limit and $\tilde{\beta}\Lambda$ should vanish as V^{-1}. On the other hand, if $T<T_c$, the condenstate occurs, ρ takes a finite value. Thus we call ρ the **condenstate density**.

For $T<T_c$, due to the condenstate, we write

$$\rho=S-\int\frac{d^dk}{(2\pi)^d}\frac{1}{e^{\tilde{\beta}\varepsilon_{\mathbf{k}}}-1}=S-I(\tilde{\beta}) \tag{48}$$

$$\tilde{F}=S-\int\frac{d^dk}{(2\pi)^d}\frac{\varepsilon_{\mathbf{k}}}{e^{\tilde{\beta}\varepsilon_{\mathbf{k}}}-1} \tag{49}$$

The results for $T=0$ is easy to derive by from the correlation function, giving

$$\lim_{\mathbf{r}\to\infty}\langle\hat{\mathbf{S}}^z(0)\cdot\hat{\mathbf{S}}^z(\mathbf{r})\rangle=0 \tag{50}$$

$$\lim_{\mathbf{r}\to\infty}\langle\hat{\mathbf{S}}^{\perp}(0)\cdot\hat{\mathbf{S}}^{\perp}(\mathbf{r})\rangle=\rho^2=m^2 \tag{51}$$

where $\perp$ denotes the $x-y$ plane. The **magnetization** $m^{\perp}=\rho$, and zero in the z direction. It is important to notice that the magnetization is actually the condensate density defined here. Similar results can be obtained for $T<T_c$ by expanding Eq. (48) and (49), with $m^{\perp}\to 0$ when $T\to T_c^-$.

Results (50) and (51) indicate that the original system has an FM order only in the $x-y$ plane at $T<T_c$, with $U(1)$ symmetry. The original SU(2) symmetry is broken at T_c, with the condensation of Schwinger bosons. In other words, the condensate region of Schwinger bosons reflects the long-range order.

IV. EXAMPLE 2: THE SU(2) FIM HEISENBERG CHAIN

A. Mean-field Hamiltonian

In this section, we move to our second example, a ferrimagnetic (FiM) Heisenberg chain, which is more non-trivial than the FM example, thus can show the efficacy of SBMFT. Though the system has changed, the procedure is similar to the FM case in Sec. III. However, remarkably, SBMFT can give even *quantitively correct* results in this model, which are consistent with many numerical calculations. I will not list all those results in this paper, and one may refer to Ref. [5]. I have consulted Prof. Wu, who is the author, and he was also surprised by this magical thing because we typically expect MFA is only qualiatively correct.

The FiM Heisenberg chain was initially designed to try explaining the ferrimagnetism by considering two types of spins $S^A = 1/2$ on sublattice A and $S^B = 1$ on sublattice B at the same time. The Hamiltonian is written as

$$\hat{H}_{\text{FiM}} = \sum_{i,\eta} \hat{\mathbf{S}}_i^A \cdot \hat{\mathbf{S}}_{i+\eta}^B \tag{52}$$

where $i = 1, \ldots, N$ and η denotes the two nearest neighbors of each site on the chain, i.e. we have $2N$ sites in total.

Introducing the SB representation (3), we rewrite

$$\hat{S}^A_{i,+} = \hat{a}^\dagger_{i,\uparrow}\hat{a}_{i,\downarrow}, \quad \hat{S}^A_{i,-} = \hat{a}^\dagger_{i,\downarrow}\hat{a}_{i,\uparrow}, \quad \hat{S}^A_{i,z} = \frac{1}{2}(\hat{a}^\dagger_{i,\uparrow}\hat{a}_{i,\uparrow} - \hat{a}^\dagger_{i,\downarrow}\hat{a}_{i,\downarrow}) \tag{53}$$

$$\hat{S}^B_{j,+} = \hat{b}^\dagger_{i,\uparrow}\hat{b}_{j,\downarrow}, \quad \hat{S}^B_{j,-} = \hat{b}^\dagger_{j,\downarrow}\hat{b}_{j,\uparrow}, \quad \hat{S}^B_{j,z} = \frac{1}{2}(\hat{b}^\dagger_{j,\uparrow}\hat{b}_{j,\uparrow} - \hat{b}^\dagger_{j,\downarrow}\hat{b}_{j,\downarrow}) \tag{54}$$

$$\tag{55}$$

Similar to the FM and AFM bond operator defined in Sec. II, we define the **FiM bond operator** as

$$\hat{D}_{i,i+\eta} = \frac{1}{2}(\hat{a}_{i\uparrow}\hat{b}_{i+\eta,\downarrow} - \hat{a}_{i,\downarrow}\hat{b}_{i+\eta,\uparrow}) \tag{56}$$

and rewrite Hamiltonian (52) as

$$\hat{H}_{\text{FiM}} = -2\sum_{i,\eta}\hat{D}^\dagger_{i,i+\eta}\hat{D}_{i,i+\eta} + \sum_{i,\eta} S_i^A S_{i+\eta}^B \tag{57}$$

At the mean-field level, we also set $\langle \hat{D}_{i,i+\eta}\rangle = D$ to be some uniform number (not necessarily real). Besides, we introduce two Lagarangian multipliers λ^A and λ^B for the two types of constraints for S^A and S^B respectively. This also comes from the considerations of lattice symmetry as what we do in Sec. III. Then we achieve the MF Hamiltonian as

$$\begin{aligned}\hat{H}_{\text{FiM-MF}} = &-\sum_{i,\eta}\left[D^*(\hat{a}_{i,\uparrow}\hat{b}_{i+\eta,\downarrow} - \hat{a}_{i,\downarrow}\hat{b}_{i+\eta,\uparrow}) + (\hat{a}^\dagger_{i,\uparrow}\hat{b}^\dagger_{i+\eta})\right] \\ &+ \lambda^A\sum_{i,\sigma}(\hat{a}^\dagger_{i,\sigma}\hat{a}_{i,\sigma} - 2S^A) + \lambda^B\sum_{j,\sigma}(\hat{b}^\dagger_{i,\sigma}\hat{b}_{i,\sigma} - 2S^B) + 2zND^*D + zNS^AS^B\end{aligned} \tag{58}$$

where $\sigma \in \{\uparrow, \downarrow\}$ and z is the coordination number. In the momentum space, we have

$$\begin{aligned}\hat{H}_{\text{FiM-MF}} = &\sum_{k,\sigma}\left[\lambda^A\hat{a}^\dagger_{k,\sigma}\hat{a}_{k,\sigma} + \lambda^B\hat{b}^\dagger_{k,\sigma}\hat{b}_{k,\sigma}\right] - \sum_{k,\sigma}\left[D^*z\gamma_k^*(\hat{a}_{k,\uparrow}\hat{b}_{-k,\downarrow} - \hat{a}_{k,\downarrow}\hat{b}_{-k,\uparrow}) + Dz\gamma_k(\hat{a}^\dagger_{k,\uparrow}\hat{b}^\dagger_{-k,\downarrow} - \hat{a}^\dagger_{k,\downarrow}\hat{b}^\dagger_{-k,\uparrow})\right] \\ &+2zND^*D - 2N(S^A\lambda^A + S^B\lambda^B) + zNS^AS^B\end{aligned} \tag{59}$$

with

$$\gamma_k := \frac{1}{z}\sum_{\eta} e^{ik\eta} \tag{60}$$

For 1D chain, $z=2$ and $\gamma_k=\cos k$. Unlike the FM case, we have to further introduce the Bogoliubov transformation to diagonalize the Hamiltonian, which is

$$\begin{bmatrix} \hat{a}_{k,\uparrow} \\ \hat{b}^{\dagger}_{-k,\downarrow} \end{bmatrix} = \begin{bmatrix} \cosh\theta_k & -\sinh\theta_k \\ -\sinh\theta_k & \cosh\theta_k \end{bmatrix} \begin{bmatrix} \hat{\beta}_{k,\uparrow} \\ \hat{a}^{\dagger}_{-k,\downarrow} \end{bmatrix} \tag{61}$$

with $\tanh 2\theta = 2|zD\gamma_k|/(\lambda^A+\lambda^B)$. Then the Hamiltonian becomes

$$\begin{aligned} \hat{H}_{\text{FiM-MF}} =& \sum_{k,\sigma}\left[E_k^{\alpha}(\hat{\alpha}^{\dagger}_{k,\sigma}\alpha_{k,\sigma}+\frac{1}{2})+E_k^{\beta}(\hat{\beta}^{\dagger}_{k,\sigma}\hat{\beta}_{k,\sigma}+\frac{1}{2})\right] \\ & +2zND^*D-2N(S^A+\frac{1}{2})\lambda^A-2N(S^B+\frac{1}{2})\lambda^B+zNS^AS^B \end{aligned} \tag{62}$$

where the two energy spectrums of α boson and β boson are

$$E^{\alpha}_{k,\sigma} = \frac{\lambda^A-\lambda^B}{2}+\sqrt{\left(\frac{\lambda^A+\lambda^B}{2}\right)^2-|zD\gamma_k|^2} \tag{63}$$

$$E^{\beta}_{k,\sigma} = -\frac{\lambda^A-\lambda^B}{2}+\sqrt{\left(\frac{\lambda^A+\lambda^B}{2}\right)^2-|zD\gamma_k|^2} \tag{64}$$

respectively. Analogously, we have to minimize the free energy to obatin the self-consistent equations. In this model, the free energy per site is given by

$$f=\frac{2}{\beta}\int_{-\pi/2}^{\pi/2}\frac{dk}{2\pi}\left\{\ln\left[2\sinh\left(\frac{\beta}{2}E_k^{\alpha}\right)\right]+\ln\left[2\sinh\left(\frac{\beta}{2}E_k^{\beta}\right)\right]\right\} \tag{65}$$

where the integral is over the first Brillouin zone. Then the self-consistent equations are

$$S^B-S^A=\int_{-\pi/2}^{\pi/2}\frac{dk}{2\pi}\left[\coth\frac{\beta}{2}E_k^{\beta}-\coth\frac{\beta}{2}E_k^{\alpha}\right] \tag{66}$$

$$S^B+S^A+1=\int_{-\pi/2}^{\pi/2}\frac{dk}{2\pi}\frac{(\lambda^A+\lambda^B)/2}{\sqrt{\left(\frac{\lambda^A+\lambda^B}{2}\right)^2-|zD\gamma_k|^2}}\left[\coth\frac{\beta}{2}E_k^{\beta}+\coth\frac{\beta}{2}E_k^{\alpha}\right] \tag{67}$$

$$\frac{2}{z}=\int_{-\pi/2}^{\pi/2}\frac{dk}{2\pi}\frac{|\gamma_k|^2}{\sqrt{\left(\frac{\lambda^A+\lambda^B}{2}\right)^2+|zD\gamma_k|^2}}\left[\coth\frac{\beta}{2}E_k^{\beta}+\coth\frac{\beta}{2}E_k^{\alpha}\right] \tag{68}$$

B. Condensation in this model

For the zero temperature or $\beta\to\infty$, at which limit the right hand side of Eq. (66) is zero unless there exists a mode whose energy is also zero to avoid the cancel of two hyperbolic cotangents. Therefore if $S^B=S^A$ (the AFM case), such a mode does not exist. Recall the calculations of the BEC in the FM model in Sec. III, such a mode extactly relates to the condensate of Schwinger Bosons. Therefore the condensation exists if and only if $S^A\neq S^B$, which accords with our knowledge that 1D AFM

Heisenberg model has no long-range order even at $T=0$. At $k=0$, when the the β branch takes energy zero, the α branch takes $2\Lambda_2$ (notice that here we use β for representing both one of the branches and the inverse temperature, but it is easy to tell). On another hand, for finite temperature, Eq. (66) always has a solution, indicating the bosons do not condense when $T>0$, or this region is disordered.

V. SBMFT TO THE SU(N) AFM HEISENBERG MODEL

As the final example, we consider the SU(N) AFM Heisenberg model

$$\hat{H}_{\text{AFM-N}} = -\frac{J}{N}\sum_{\langle ij\rangle}(\hat{A}^\dagger_{ij}\hat{A}_{ij} - NS^2) \tag{69}$$

using the functional integral approach, where

$$\hat{A}_{ij} = \sum_{\mu=1}^{N}\hat{b}_{i\mu}\hat{b}_{j\mu} \tag{70}$$

We will not go over the details of this approach, only presenting some main results. For $\tau\in[0,\beta]$ to be the imaginary time, two extra fields are introduced in this approach: (i) At each site, a single real field $\lambda_i(\tau)$ which ensures the occupancy constraint; (ii) On each link, a complex Hubbard-Stratonovich field $Q_{ij}(\tau)$ which can decouple the interaction.

At the mean-field level, we assume the two fields to be static, i.e. independent of τ, denoted by λ_i and Q_{ij} respectively. This is similar to the direct mean-field approach in Sec. III and Sec. IV, where we fix the mean value of the bond operator and consider some uniform parameter as the Lagarangian multiplier.

The corresponding mean-field Hamiltonian is written as

$$\hat{H}^{\text{MF}}_{\text{AFM-N}} = \frac{N}{J}\sum_{\langle ij\rangle}|Q_{ij}|^2 + \sum_{\langle ij\rangle}(Q_{ij}\hat{A}^\dagger_{ij} + Q^*_{ij}\hat{A}_{ij}) + \sum_{i\mu}\lambda_i(\hat{b}^\dagger_{i\mu}\hat{b}_{i\mu} - \hat{n}_s) + \frac{1}{\sqrt{VN}}\sum_{i\mu}(\phi^*_{i\mu}\hat{b}_{i\mu} + \phi_{i\mu}\hat{b}^\dagger_{i\mu}) \tag{71}$$

where $\hat{n}_s := \sum_{\mu=1}^{N}\hat{b}^\dagger_{i\mu}\hat{b}_{i\mu}$ denotes the total number of Schwinger bosons, andthe scalar field $\phi_{i\mu}$ linearly coupled to the Schwinger bosons is defined by

$$\frac{\partial F}{\partial\phi^*_{i\mu}} = \frac{\langle\hat{b}_{i\mu}\rangle}{\sqrt{VN}} \tag{72}$$

Recall that the ground stated of BEC is described by a coherent state $|\psi\rangle$, defined by

$$|\psi\rangle := e^{\psi\hat{a}^\dagger}|\text{vacuum}\rangle = \sum_{n=0}^{\infty}\frac{\psi^n(\hat{a}^\dagger)^n}{n!}|\text{vacuum}\rangle = \sum_{n=0}^{\infty}\frac{\psi^n}{\sqrt{n!}}|n\rangle \tag{73}$$

for $\hat{a}^\dagger$ ($\hat{a}$) to be the corresponding bosonic creation (annihilation) operator. satisfying

$$\hat{a}|\psi\rangle = \psi|\psi\rangle, \quad \langle\psi|\hat{a}^\dagger = \langle\psi|\psi^* \tag{74}$$

Therefore both $\langle \hat{a}^\dagger \rangle$ and $\langle \hat{a} \rangle$ can be used as the order parameter since they are non-zero in the BEC phase, zero otherwise. In our situation, for the definition (72), it is helpful to replace $\phi_{i\mu}$ with $\langle \hat{b}_{i\mu} \rangle$, which offers more convenience when discussing the condensate. Therefore we can introduce the Legendre transformation

$$G = F - \sum_{i\mu} (\phi_{i\mu} \beta_{i\mu}^* + \phi_{i\mu}^* \beta_{i\mu}) \tag{75}$$

where

$$\beta_{i\mu} := \frac{\langle b_{i\mu} \rangle}{\sqrt{VN}} \tag{76}$$

and both G and F in Eq. (75) are used to denote the free energy.

With some algebras, the final form of the free energy per site and per flavor is

$$g = \frac{G}{VN} = \frac{z}{2J}|Q|^2 - (\kappa + \frac{p}{2})\lambda + p \int_{\mathrm{BZ}} \frac{d^d k}{(2\pi)^d} \left[\frac{1}{2}\omega_{\mathbf{k}} + T \ln(1 - e^{-\omega_k/T}) \right] + E_{\mathrm{con}} \tag{77}$$

where $\kappa = n_s/N$ indicating the average boson number for each flavor, and the condensation energy is given by

$$E_{\mathrm{con}} = \lambda \sum_{\mathbf{k}\mu} |\beta_{\mathbf{k}\mu}|^2 + \frac{1}{2} z \sum_{\mathbf{k}\mu\mu'} (Q\gamma_{\mathbf{k}} \beta_{\mathbf{k}\mu}^* \beta_{-\mathbf{k}\mu'}^* + Q^* \gamma_{\mathbf{k}}^* \beta_{\mathbf{k}\mu} \beta_{-\mathbf{k}\mu'}) \tag{78}$$

with dispersion relation

$$\omega_{\mathbf{k}} = \sqrt{\lambda^2 - |zQ\gamma_{\mathbf{k}}|^2} \tag{79}$$

Similarly, by minimizing the free energy g, we can obtain the following self-consistent equations

$$\kappa + \frac{1}{2} = \int_{\mathrm{BZ}} \frac{d^d k}{(2\pi)^d} \frac{\lambda}{\omega_{\mathbf{k}}} \left[n_{\mathbf{k}}(T) + \frac{1}{2} \right] + \sum_{\mathbf{k}\alpha} |\beta_{\mathbf{k}\alpha}|^2 \tag{80}$$

$$\frac{z}{J}|Q|^2 = p \int_{\mathrm{BZ}} \frac{d^d k}{(2\pi)^d} \frac{|zQ\gamma_{\mathbf{k}}|}{\omega_{\mathbf{k}}} \left[n_{\mathbf{k}}(T) + \frac{1}{2} \right] + \lambda \sum_{\mathbf{k}\mu} |\beta_{\mathbf{k}\mu}|^2 \tag{81}$$

$$0 = \lambda \beta_{\mathbf{k}\mu} + zQ\gamma_{\mathbf{k}} \sum_{\mu'} \delta_{\mu\mu'} \beta_{-\mathbf{k}\mu'}^* \tag{82}$$

where

$$n_{\mathbf{k}} = \frac{1}{e^{\omega_{\mathbf{k}}/T} - 1} \tag{83}$$

Assume the condensate only occurs at some unique mode $\mathbf{k}$, then with some algebras, we will find $\omega_{\mathbf{k}} = 0$, ensuring the gapless specturm when the BEC occurs. For $T = 0$, there exists some critical value of κ, denoted as κ_c, above which the BEC occurs. Such a value can be derived by

$$\kappa_c = \frac{1}{2} \int_{\mathrm{BZ}} \frac{d^d k}{(2\pi)^d} \frac{1}{\sqrt{1 - |\gamma_{\mathbf{k}}|^2}} - \frac{1}{2} \tag{84}$$

Eq. (84) has no solution for $d = 1$, meaning a condensate will never happen. This also accords with our knowledge that the SU(2) Heisenberg chain has no long-range order even at $T = 0$. For the two dimensional square lattice, there actually exists a solution of $\kappa_c = 0.19$. Still we consider the SU(2) case, where even the minimum value of $S = 1/2$ can be Néel ordered as $\kappa = 2S/2 = S$.

VI. SUMMARY

The Schwinger boson representation offers an alternative way to study the Heisenberg model. It is not limited to ferrimagnetic or antiferromagnetic, frustrated or non-frustrated, single- or mixed-spins systems, and can be applied to both ordered and disordered region in any dimensions. Another advantage of this representation is that it can be easily generalized to the SU(N) Heisenberg model, which, based on the large-N expansion, can help understand the physical SU(2) case.

Combined with mean-field theory, many important qualitatively even quantitively properties can be derived. We have introduced three examples in this paper, including the SU(2) ferromagnetic model, the SU(2) ferrimagnetic chain, and the SU(N) antiferromagnetic model. The long-range order of the spin model is reflected by the condensate of corresponding Schwinger (Bogoliubov) bosons, which can be used to further characterize the thermodynamic properties of the models.

In addition, more applications of the Schwinger boson representation can be explored. For example, it can be used to construct the valence bond states [M. Raykin and A. Auerbach, Phys. Rev. B 47, 5118 (1993)], even to calculate quantum entanglement [H. Katsura et al. Phys. Rev. B 76, 012401 (2007)].

ACKNOWLEDGEMENT

This paper was completed under the guidance of Prof. Congjun Wu, and I thank him for his great kindness and patience.

References

(sorted by year)

[1] A. Auerbach and D.P. Arovas, Phys. Rev. Lett. 61, 617 (1988). *Sec. II, Sec V*

[2] D.P. Arovas and A. Auerbach, Phys. Rev. B 40, 791 (1989). *Sec. II, Sec V*

[3] A. Auerbach, SU(N) Heisenberg Models. In: Interacting Electrons and Quantum Magnetism. Graduate Texts in Contemporary Physics. Springer, New York, NY (1994). *Sec. II, Sec V*

[4] S. Sarker, C. Jayaprakash, H.R. Krishnamurthy, and M. Ma, Phys. Rev. B 40, 5028 (1989). *Sec. III*

[5] C. Wu, B. Chen, X. Dai, Y. Yu, and Z.B. Su, Phys. Rev. B 60, 1057 (1999). *Sec. IV*

[6] A. Auerbach, D.P. Arovas, Schwinger Bosons Approaches to Quantum Antiferromagnetism. In: Lacroix, C., Mendels, P., Mila, F. (eds) Introduction to Frustrated Magnetism. Springer Series in Solid-State Sciences, vol 164. Springer, Berlin, Heidelberg (2011). *Sec. II, Sec V*