Real-space pairing in an extended t-J model

Jakub Jędrak

Rozprawa doktorska

Promotor: Prof. dr hab. Józef Spałek

Uniwersytet Jagielloński
Instytut Fizyki im. Mariana Smoluchowskiego
Zakład Teorii Materii Skondensowanej i Nanofizyki

Kraków, kwiecień 2011

Abstract

High-temperature superconductivity in copper oxides (cuprates, e.g. $La_{2-x}Sr_xCuO_4$, $YBa_2Cu_3O_{6+\delta}$ or $Bi_2Sr_2CaCu_2O_{8+\delta}$) remains among the most spectacular phenomena in condensed matter physics. Since its discovery in 1986, an enormous number ($\sim 10^5$) of papers on the subject have appeared. So far, there is no single, commonly accepted theory of high-temperature superconductivity. However, it is widely believed that a basic description of this phenomenon can be provided by a single-band Hubbard model or its derivative, the t-J model. The latter model is regarded as a minimal microscopic model, capable of describing the essential aspects of the complex physics of the cuprates.

Unfortunately, in general case, neither of those two models can be solved exactly and therefore various approximate methods are used. Among them, the so-called *mean-field methods* provide a simple, yet fairly reasonable description of the cuprates. In particular, some of the main qualitative features of the phase diagram and the essential features of electronic spectrum are roughly reproduced.

A standard mean-field approach to the t-J model, known under the name of *renormalized mean-field theory* (RMFT) goes beyond the Hartree-Fock approximation. Consequently, its fully consistent treatment requires a novel theoretical approach. This has been our original motivation to develop a general approach to the mean-field models, which is based on the *maximum entropy* (MaxEnt) *principle*. The method is presented in detail in Part II of this Thesis, and in Part III it is applied to study RMFT of the t-J model. First, we compare the results obtained within our formalism with those of the frequently used non-variational approach based entirely on the self-consistent equations. Also, various versions of RMFT are compared, and the most satisfactory of them is selected. This optimal version is subsequently used to study different versions of the original t-J Hamiltonian. As a result, upper critical concentration and doping dependence of the selected physical quantities (e.g. the superconducting gap and the Fermi velocity) is determined at low temperatures and in the absence of external magnetic field. We compare our findings both with theoretical results obtained from the Variational Monte Carlo (VMC) methods, as well as with the experimental data for selected cuprates. We show that the version of RMFT approach formulated in this Thesis provides a reasonable qualitative and in some cases semiquantitative rationalization of the principal characteristics of the hole-doped high-temperature superconductors at the optimal doping and in the overdoped regime.

Possible extensions of the proposed analysis are mentioned at the end.

Keywords: *High-T_c superconductivity, cuprates, phase diagram for high-T_c compounds, strongly correlated fermions, resonating valence-bond (RVB) state, t-J model, Gutzwiller projection, Gutzwiller approximation, Maximum Entropy (MaxEnt) principle, mean field theory.*

Streszczenie

Nadprzewodnictwo wysokotemperaturowe w tlenkach miedzi (krótko: miedzianach, np. $La_{2-x}Sr_xCuO_4$, $YBa_2Cu_3O_{6+\delta}$ lub $Bi_2Sr_2CaCu_2O_{8+\delta}$) pozostaje jednym z najbardziej spektakularnych zjawisk w fizyce materii skondensowanej. Od jego odkrycia w roku 1986, ukazała się ogromna liczba ($\sim 10^5$) prac poświęconych tej tematyce. Do tej pory nie istnieje jedna, powszechnie akceptowana teoria nadprzewodnictwa wysokotemperaturowego. Niemniej jednak, uważa się prawie powszechnie, iż prawidłowy opis tego zjawiska można uzyskać w ramach jednopasmowego modelu Hubbarda lub wywodzącego się zeń modelu t-J. Ten ostatni jest uważany także za minimalny model mikroskopowy, zdolny opisać istotne aspekty struktury stanów elektronowych i związanej z nią złożonej fizyki związków na bazie tlenku miedzi.

Niestety, w ogólnym przypadku, żadnego z wyżej wymienionych modeli nie można rozwiązać w sposób ścisły, i dlatego też używa się różnych metod przybliżonych. Miedzy innymi, tzw. *metody pola średniego* stanowią rozsądny kompromis pomiędzy prostotą opisu a jego dokładnością. W szczególności, z grubsza odtworzone zostają główne cechy diagramu fazowego, a także struktura elektronowa nadprzewodników na bazie tlenku miedzi.

Standardowa metoda typu pola średniego dla modelu t-J, znana pod nazwą *zrenormalizowanej teorii pola średniego* (ang. *renormalized mean-field theory*, RMFT), wykracza poza przybliżenie Hartree-Focka. Z tego powodu, w pełni wewnętrznie spójne potraktowanie zrenormalizowanej teorii pola średniego wymaga nowego podejścia teoretycznego.

Idea takiego podejścia stanowiła w tej rozprawie motywację do rozwinięcia ogólnego podejścia do metod typu pola średniego, podejścia opartego na *zasadzie maksimum entropii*, (MaxEnt) (ang. *maximum entropy principle*). Podejście to jest szczegółowo przedstawione w części II rozprawy, zaś w części III zostaje zastosowane do badania zrenormalizowanej teorii pola średniego dla modelu t-J. W części III zaczynamy od porównania wyników otrzymanych w ramach naszego formalizmu z wynikami często używanego podejścia niewariacyjnego, opartego w całości na tzw. równaniach samouzgodnionych Bogoliubowa-de Gennesa. Porównane zostają także różne wersje RMFT, a następnie jedna z nich, o najbardziej z punktu widzenia eksperymentu zadowalających własnościach, zastosowana jest do badania różnych wersji pełnego Hamiltonianu t-J. W rezultacie, w temperaturach bliskich zera bezwzględnego i przy braku zewnętrznego pola magnetycznego, wyznaczona zostaje górna koncentracja krytyczna i zależności wybranych własności fizycznych (np. przerwy nadprzewodzącej oraz prędkości Fermiego) od stopnia domieszkowania układu. Nasze wyniki teoretyczne są następnie porównane z wynikami podejścia typu 'Variational Monte Carlo' (VMC), a także z danymi doświadczalnymi dla wybranych miedzianów. Pokazujemy, iż wersja RMFT sformułowana w tej rozprawie prowadzi do rozsądnego opisu głównych cech wysokotemperaturowych nadprzewodników miedziowych domieszkowanych dziurowo, oraz jakościowej, a w pewnych przypadkach półilościowej, zgodności z doświadczeniem, tak przy domieszkowaniu optymalnym, jak i większym od optymalnego.

Możliwe uogólnienia zaproponowango tu podejścia są przedstawione na końcu rozprawy. Poza tym, w całej rozprawie staramy się omówić krytycznie zasadnicze cechy opisywanego podejścia.

Contents

Podziękowania

Profesorowi Józefowi Spałkowi, promotorowi niniejszej rozprawy, jestem bardzo wdzięczny za zaproponowanie tematyki badań, za cenne dyskusje oraz za wszelkie otrzymane uwagi, zarówno merytoryczne, jak i językowe. Dziękuję Mu także za wsparcie, wyrozumiałość i cierpliwość, jakimi darzy mnie w ciągu wielu lat naszej współpracy.

Profesorowi Krzysztofowi Rościszewskiemu bardzo dziękuję za cenne uwagi i wskazówki, które otrzymałem od Niego na wczesnym etapie pracy nad zagadnieniami przedstawionymi w niniejszej rozprawie.

Chciałbym podziękować Wszystkim, którzy w mniejszym lub większym stopniu przyczynili się do powstania niniejszej pracy swoimi radami, zachętą, uwagami i dyskusjami, jak również wszelakiego rodzaju pomocą: Marcinowi Abramowi, doktorowi Andrzejowi Biborskiemu, doktorowi Michałowi Hellerowi, Oldze Howczak, Janowi Kaczmarczykowi, doktorowi Andrzejowi Kapanowskiemu, Małgorzacie Kaliszan, Michałowi Kłosowi, Magdalenie Kozłowskiej, Jaromirowi Krzyszczakowi, doktorowi Romanowi Marcinkowi, doktorowi Marcinowi Raczkowskiemu, doktor Joannie Sapetowej, Zygmuntowi Starypanowi, Katarzynie Targońskiej, doktorowi Krzysztofowi Wohlfeldowi oraz Marcinowi Wysokińskiemu.

Pragnę także gorąco podziękować Profesorowi Ehudowi Altmanowi i Jego Współpracownikom, w szczególności doktor Lilach Goren, za życzliwą gościnę, cenne dyskusje i krytyczne uwagi jakie otrzymałem w trakcie mojego krótkiego pobytu w Instytucie im. Chaima Weizmanna w Rehovot.

Jestem bardzo wdzięczny Profesorowi Florianowi Gebhardowi i doktorowi Jörgowi Bünemannowi, za uwagi i dyskusje, które miały miejsce podczas konferencji 'Korrelationstage 2011' w Instytucie im. Maksa Plancka w Dreźnie.

Moja ogromna wdzięczność należy sie rownież Autorom biblioteki GSL (Gnu Scientific Library), w oparciu o którą wykonane zostały wszystkie obliczenia numeryczne przedstawione w niniejszej rozprawie.

Niniejsza rozprawa była częściowo finansowana z grantu (N N 202 128 736) Ministerstwa Nauki i Szkolnictwa Wyższego Rzeczpospolitej Polskiej.

List of frequently used abbreviations

MF	mean-field
MaxEnt	maximum entropy
GC	grand canonical
HF	Hartree-Fock
RVB	resonating valence bond
BCS	Bardeen-Cooper-Schrieffer
GWF	Gutzwiller wave function
GA	Gutzwiller approximation
RMFT	renormalized mean-field theory
RS	renormalization scheme
SBMFT	slave-boson mean-field theory
SC	superconducting
AF	antiferromagnetic
PG	pseudogap
PI	Pomeranchuk instability
SF	staggered flux
ARPES	angle resolved photoemission spectroscopy
FS	Fermi surface
FFLO	Fulde-Ferrell-Larkin-Ovchinnikov
var	variational
s-c	self-consistent

List of frequently used symbols

$A_1, A_2, \ldots, A_M$	mean fields
$\vec{A} = (A_1, A_2, \ldots, A_M)$	vector of mean fields
$\hat{H}(\vec{A}) \equiv \hat{H}$	original mean-field Hamiltonian
$\langle \hat{A} \rangle$	average value of $\hat{A}$ operator
$\mathcal{D}_{\mathcal{A}}$	domain of $\hat{H}(\vec{A})$
D_S	spatial dimension of the lattice
D_H	dimension of the Hilbert space
$\hat{N}$	particle number operator
Λ	number of lattice sites
$\hat{H}_e$	'exact', i.e., non-mean-field Hamiltonian
$\hat{\rho}_e$	'exact', i.e., non-mean-field density operator
μ	chemical potential
T	temperature
$\beta = 1/k_B T$	inverse temperature
$S_{vN}(\hat{\rho})$	von Neumann entropy for density operator $\hat{\rho}$
$\mathcal{S}_e$	entropy functional
$\mathcal{S}_\lambda$	mean-field entropy functional supplemented with the self-consistency-preserving constraints.
$\sum_{s=1}^{M} \lambda_s (\mathrm{Tr}[\hat{\rho}_\lambda \hat{A}_s] - A_s)$	self-consistency preserving constraints
$\lambda_1, \lambda_2, \ldots, \lambda_M$	Lagrange multipliers
$\vec{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_M)$	vector of Lagrange multipliers
$\hat{H}_\lambda = \hat{H} - \sum_{s=1}^{M} \lambda_s (\hat{A}_s - A_s)$	mean-field Hamiltonian supplemented with the constraint terms
$\hat{K}_\lambda \equiv \hat{H}_\lambda - \mu \hat{N}$	mean-field grand Hamiltonian corresponding to $\hat{H}_\lambda$
$\hat{\rho}_\lambda = \mathcal{Z}_\lambda^{-1} \exp\big(-\beta \hat{K}_\lambda\big)$	mean-field density operator corresponding to $\hat{K}_\lambda$
$\mathcal{Z}_\lambda^{-1} = \mathrm{Tr}[\exp\big(-\beta \hat{K}_\lambda\big)]$	mean-field partition function
$p_i,\ q_i$	probability of i-th microstate
$b_1, b_2, \ldots, b_P$	variational parameters of the non-mean-field character
$(b_1, b_2, \ldots, b_P) \equiv \vec{b}$	vector of b_l parameters
$\mathcal{F}(\vec{A}, \vec{\lambda}, \vec{b}) \equiv -\beta^{-1} \ln \mathcal{Z}_\lambda(\vec{A}, \vec{\lambda}, \vec{b})$	generalized grand potential
$\vec{A}_0,\ \vec{\lambda}_0,\ \vec{b}_0$	optimal (equilibrium) values of $\vec{A}$, $\vec{\lambda}$, and $\vec{b}$, respectively
$\hat{K}_{\lambda 0} = \hat{K}_\lambda(\vec{A}_0, \vec{\lambda}_0, \vec{b}_0)$	equilibrium mean-field grand Hamiltonian
$\hat{\rho}_{\lambda 0} = \mathcal{Z}_{\lambda 0}^{-1} \exp(-\beta \hat{K}_{\lambda 0})$	equilibrium mean-field density operator
$\vec{A}_{sc}^{(0)}$	Optimal solution of the self-consistent equations obtained within the non-variational (Bogoliubov-de Gennes) approach.
$\vec{\lambda}(\vec{A})$	optimal solution of the self-consistent equations obtained within the present approach
$\hat{H}_z(\vec{A}) \equiv \hat{H}_\lambda(\vec{A}, \vec{\lambda}(\vec{A}))$	self-consistent mean-field Hamiltonian
$\hat{\rho}_z(\vec{A}) = \hat{\rho}_\lambda(\vec{A}, \vec{\lambda}(\vec{A}))$	self-consistent mean-field density operator
$\mathcal{F}_z(\vec{A}) \equiv \mathcal{F}(\vec{A}, \vec{\lambda}(\vec{A}))$	self-consistent grand potential (Landau potential)
$\Omega(T, V, \mu, \vec{h})$	thermodynamic grand potential
$F(T, V, N, \vec{h}) = \Omega + \mu N$	free energy
$W(\vec{A}) = \langle \hat{H}(\vec{A}) \rangle$	average value of the mean-field Hamiltonian

$\vec{R}_i$	position vector of the i-th lattice site
$\lvert\vec{R}_i - \vec{R}_j\rvert = d(i,j)$	distance between i-th and j-th lattice sites
$\hat{H}_{tJ}$	t-J Hamiltonian
$\hat{H}_t$	kinetic energy part of the t-J Hamiltonian
$\hat{H}_J$	exchange part of the t-J Hamiltonian
$\hat{H}_3$	three-site term part of the t-J Hamiltonian
$\hat{H}_{tU}$	Hubbard Hamiltonian
$\hat{H}_{tJU}$	t-J-U Hamiltonian
t_{ij}	hopping integral between lattice sites labeled by $\vec{R}_i$ and $\vec{R}_j$
J_{ij}	exchange integral between lattice sites labeled by $\vec{R}_i$ and $\vec{R}_j$
$\hat{P} = \prod_i (1 - \hat{n}_{i\uparrow}\hat{n}_{i\downarrow})$	Gutzwiller projection operator
$\lvert\mathrm{BCS}\rangle$	Bardeen-Cooper-Schrieffer (BCS)-type state
$\lvert\mathrm{RVB}\rangle = \hat{P}\lvert\mathrm{BCS}\rangle$	resonating valence bond state
$\lvert\Psi\rangle = \hat{P}_C\lvert\Psi_0\rangle$	correlated trial state
$\hat{P}_C$	correlator
$\lvert\Psi_0\rangle$	eigenstate of a single-particle Hamiltonian
$\langle\hat{\mathcal{O}}\rangle_C \equiv \langle\Psi\lvert\hat{\mathcal{O}}\rvert\Psi\rangle / \langle\Psi\vert\Psi\rangle$	correlated average of operator $\hat{\mathcal{O}}$
$\langle\hat{\mathcal{O}}\rangle$	uncorrelated, (i.e., computed using $\lvert\Psi_0\rangle$) average of $\hat{\mathcal{O}}$
$g^{\mathcal{O}}$	renormalization factor for operator $\hat{\mathcal{O}}$.
g^t_{ij}	renormalization factor for the kinetic energy
g^J_{ij}	renormalization factor for the spin exchange interaction
$\hat{\rho}_0$	grand canonical single-particle mixed state
$\hat{\rho}_{\mathrm{C}} = \hat{P}_C\hat{\rho}_0\hat{P}_C$	correlated mixed state
$\hat{H}_R$	RMFT Hamiltonian
$\hat{H}_{R\lambda}$	RMFT Hamiltonian supplemented with the constraint terms
$\hat{H}_R^{(\sim)} = W(\chi_{ij\sigma}, \Delta_{ij}, n_{i\sigma})\hat{\mathbf{1}}_{D_H}$	alternative form of $\hat{H}_R$
$\hat{H}_{R\lambda}^{(\sim)}$	alternative form of $\hat{H}_{R\lambda}$
$W(\chi_{ij\sigma}, \Delta_{ij}, n_{i\sigma}) = \langle\hat{H}_R\rangle = \langle\hat{H}_e\rangle_C^{app}$	exact expectation value of $\hat{H}_R$ (approximate expectation value of $\hat{H}_e$)
$\chi_{ij\sigma} \equiv \langle c^\dagger_{i\sigma} c_{j\sigma}\rangle$	hopping amplitude (bond order parameter)
$\Delta_{ij} \equiv \langle c_{i\bar{\sigma}} c_{j\sigma}\rangle = \langle c_{j\bar{\sigma}} c_{i\sigma}\rangle$	superconducting gap parameter
$\Delta_{Cij} \equiv \langle\hat{\Delta}_{ij}\rangle_C$	superconducting order parameter
$c^\dagger_{i\sigma}$ $(c_{i\sigma})$	creation (annihilation) operator for electron with spin $\sigma = \pm$ on the site labeled by $\vec{R}_i$
$n = \langle\hat{N}\rangle/\Lambda = N/\Lambda$	average number of electrons per lattice site
$x = 1 - n$	hole doping
$\mathbf{k}$	quasimomentum
$\xi_{\mathbf{k}}$	quasiparticle energy in the normal state
$D_{\mathbf{k}}$	superconducting gap
$E_{\mathbf{k}}$	quasiparticle energy
v_F	Fermi velocity

Part I
Introduction

1 High-temperature superconductivity of cuprate compounds and its basic theoretical models

1.1 General characteristics

High-temperature (high-T_c) superconductivity, in particular that of the cuprate compounds (cuprates), is one of the most puzzling and challenging subjects in condensed matter physics [1, 2, 4, 5, 6, 7, 8]. Since its discovery by Bednorz and Müller in 1986 [9], there is still a large interest in this field. It is partly due to potentially revolutionary technological applications - for most of high-T_c cuprate compounds the critical temperature (T_c) exceeds 77K, i.e., the boiling temperature of liquid nitrogen. From the point of view of a physicist, cuprates are exciting due to the complex structure and unusual properties of those materials.[1]

We should mention right at the beginning that it is not our aim here to analyze in detail the large number of the existing experimental data for the cuprates. Rather, we invoke only the basic facts and focus on properties, which can be described or even predicted by simple theoretical models and methods we use.

A number of high-T_c cuprate compounds have been discovered. The most notable are $La_{2-x}Sr_xCuO_4$ (LSCO), with the maximal critical temperature T_c (which however depends on the hole doping x) equal 36K, $YBa_2Cu_3O_{6+\delta}$ (YBCO) with $T_c \leq 91$K, and $Bi_2Sr_2CaCu_2O_{8+\delta}$ (BSCCO or more precisely, Bi2212) with $T_c \leq 89$K [2]. As suggested by their chemical formulas, all cuprate compounds have one or more CuO_2 plains, separated by atoms of other elements. All exhibit strong tetragonal anisotropy in the c-axis direction, and their quasi-two dimensional structure seems to be responsible for many of the essential properties of those materials. Additionally, for some high-T_c compounds, a weaker in-plane anisotropy between a and b axes may appear (orthorhombic structure). The doping-temperature (x-T) phase diagram of all *hole-doped*[2] high-T_c compounds (cf. Fig. 1) have a similar structure [1, 2, 4, 8]. Upon the hole doping, with the hole concentration $x \gtrsim 0.02-0.05$, a generic antiferromagnetic (AF) Mott insulating state of the undoped parent compound [10, 11] eventually transforms (for $x \approx 0.05$) into a superconducting (SC) state of a $d_{x^2-y^2}$ (d-wave) symmetry [12]. Still, even in absence of the long-range antiferromagnetic (AF) order, the antiferromagnetic correlations seem to be present in the SC state. The latter, in turn, after reaching a maximal transition temperature at $x \approx 0.15-0.2$, disappears at the upper critical concentration $x_c \approx 0.25-0.35$, depending on the compound [13, 14]. In the overdoped regime $x \gtrsim 0.15-0.2$ the system evolves gradually from a non-Fermi liquid into a quantum liquid that can be regarded as an unconventional Fermi liquid [15].

The region of the phase diagram where superconductivity appears is called a 'dome' due to its characteristic shape. For some cuprate compounds, antiferromagnetic and superconducting

[1] Apart from the cuprates, the class of high-temperature superconductors encompasses also the recently discovered iron-based superconductors, like pnictides, e.g. $Ba_{1-x}K_xFe_2As_2$ or oxypnictides, e.g. $GdFeAsO_{0.85}$. It should be also noted, that organic superconductors, e.g. $(TMTSF)_2PF_6$, although having $T_c \sim 1-10$K, share many properties with both copper and iron superconductors [4].

[2] There exist also electron-doped high-T_c cuprate compounds, e.g. $Nd_{2-x}Ce_xCuO_4$ ($T_c = 23$K). The generic x-T phase diagram of electron-doped compounds exhibits remarkable quantitative differences as compared to that of its hole-doped counterpart [2, 3]. Although here we concentrate on the hole-doped case, note that essentially the same theoretical methods which are developed in this Thesis may also be used to study the electron-doped compounds.

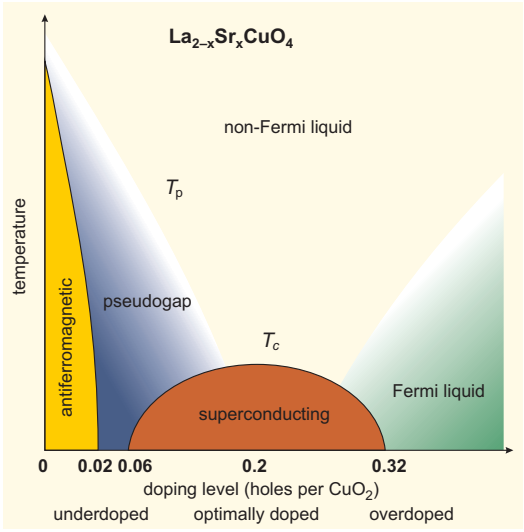

Figure 1: Schematic hole doping (x) - temperature (T) phase diagram of $La_{2-x}Sr_xCuO_4$, taken as an example of a generic hole-doped cuprate superconductor. The vertical solid line marks qualitatively the division into underdoped and the overdoped regimes.

orders occur simultaneously, i.e., we have the AF-SC phase coexistence. For others, the AF and SC regions of the phase diagram are separated by a disordered ('glassy') state.

Finally, one of the most intriguing features of the cuprates is the existence of an unconventional normal state, called 'pseudogap' (PG) or 'spin gap' [6, 8]. Pseudogap phase is visible in various experiments [16, 17, 18] above the superconducting dome in the underdoped regime. In this phase, the gaped behavior in the temperature dependence of the NMR relaxation rate is observed [8]. Also, both NMR and ARPES experiments show that magnetic excitations are suppressed in the temperature range $T_c < T < T^*$, and that the energy gap is gradually formed in one-particle excitations below T^*. The pseudogap behavior is often interpreted as an offset of the pre-formed pairs with the $d_{x^2-y^2}$ -like quasimomentum ($\mathbf{k}$) -dependence as in the SC phase [8].

1.1.1 Microscopic models of electronic states

In order to provide a theoretical description of the cuprate superconductivity, the Hubbard model is often invoked. Both the simplest, single-band form [19, 20, 21], as well as the more realistic three band (d-p model, see [8] and References therein) are used. The former model results from ascribing a passive role to the electrons on p_x and p_y oxygen orbitals and retaining only the dynamics of electrons on the copper $3d_{x^2-y^2}$ orbitals. In the strong-coupling limit (i.e., with the Coulomb interaction dominant over the kinetic energy of the electrons), the single-band Hubbard model can be transformed into the t-J model [8, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31], which is often regarded as a minimal, purely electronic microscopic model of high-T_c superconductivity. Unfortunately, as for the most of the realistic models of interacting electrons, exact solutions of the t-J model are limited to very special choice of the model parameters or to very small clusters [32]. Consequently, approximate methods of various kinds must be invoked.

1.1.2 Resonating valence bond (RVB) state

A theoretical concept which also seems to be important for the description of high-T_c superconductivity is that of the resonating valence bond (RVB) state [33]. As mentioned in the latter Reference, the notion of resonating valence bonds has been introduced by Pauling in the early

years of quantum chemistry [34, 35], e.g. to explain the nature of the electronic structure of benzene. In condensed matter physics, the RVB state has been originally used as a possible variational ground state of the Heisenberg Hamiltonian on frustrated lattices [36, 37]. Later, it has been proposed by Anderson [38] (cf. also Refs. [39, 40]) as a candidate for the ground state of a generic strongly-correlated two-dimensional superconductor.

RVB state is as a coherent superposition of electron spin-singlets residing on different pair of sites (bonds); hence the name. Due to the lack of the long-range magnetic order, it is an example of a *spin-liquid* state. On the other hand, RVB state is a Bardeen-Cooper-Schrieffer (BCS) state [41, 42, 43, 44] with doubly occupied configurations in the real space being excluded via the so-called *Gutzwiller projection* [8, 20, 33]. In other words, RVB state may be expected to play a similar role for a description of the high-T_c superconductors, as its uncorrelated counterpart, i.e., the BCS state plays in the theory of conventional superconductivity. The original RVB state may be generalized in several ways, e.g. by including the correlation effects in a more sophisticated manner or by implementing more complex patterns of the symmetry breaking [8, 33].

In one dimension ($D_S = 1$), at the half-filling ($x = 0$), a chain of singlets has lower energy than the Neel antiferromagnetic state. For $D_S = 2$ this is no longer the case; simple 'static' singlet covering yields the energy higher than the antiferromagnetic state, nevertheless, the true RVB state remains competitive to the Neel-ordered state [33]. Consequently, in two dimensions, the RVB state seems to be a reasonable variational *Ansatz* for the ground state of t-J and related models. On the technical level, this idea may be realized in two different ways. First, the expectation value of any operator (in particular, of the t-J Hamiltonian) in the RVB state may be computed by means of the Variational Monte Carlo (VMC) method [45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 56, 57]. Alternatively, RVB picture may be implemented by using an appropriate form of the *mean-field* (MF) approach.

1.2 Mean-field description of high-T_c superconductors

In this Thesis, we focus on a particular mean-field (MF) approach to the t-J model, known under the name of the *renormalized mean-field theory* (RMFT) [8, 33, 58]. RMFT is an effect of applying Gutzwiller approximation (GA) [20, 59, 60], originally devised for the Hubbard model, to the t-J model. The resulting single-particle picture is widely used due to its 'clarity and directness' [39, 61]. Moreover, RMFT is capable of reproducing the basic qualitative, and even some quantitative features of phase diagram of the cuprates [8, 33, 39]. This may be quite surprising, because in contrast to conventional superconductors, such as Al, Sn or Pb, well described by the BCS theory of a mean-field character, for the cuprates the MF approximation seems to be less adequate, for the following reason. Namely, conventional superconductors are characterized by a large coherence length. Therefore, the average distance between Cooper pairs is much smaller then the pair size (~ 1000Å), and each pair is immersed in, and interacts with many other pairs. This is the physical cause of the striking success of the Hartree-Fock approximation and BCS theory in those systems. On the other hand, high-T_c cuprates are characterized by a small coherence length, the average pair size is in the range $\sim 10 - 30$Å [7], i.e., it is only moderately greater then the average distance between electrons in the CuO_2 plane. Therefore, we cannot invoke the same argument for the validity of the MF approach as in the case of BCS superconductors, and MF treatment of the cuprates requires an alternative justification.

The basic question is then whether we can regard RMFT as a satisfactory theoretical description of high-T_c compounds, despite its simplistic nature and apparent shortages. This point of view has been advocated strongly by Anderson and coworkers [39, 40, 61], and RMFT has been, and still is, widely used in studies of the cuprates, cf. e.g. [33, 40, 58, 60, 61, 62,

63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84]. However, it has been also pointed out that RMFT can be placed in the Fermi-liquid paradigm (cf. e.g. [59, 85]), and, as such, is not expected to provide a correct description of the whole phase diagram, but may work well around and above the optimal doping [86]. Intuitively, with the increasing doping, charge carriers (holes) become more mobile, and single particle description works better. On the other hand, with the decreasing doping, charge fluctuations are smaller, and eventually vanish in the Mott-insulating limit ($x = 0$). Consequently, one may expect large phase fluctuations in the wave function describing the superconducting ground state. However, phase fluctuations are not included in the standard RMFT approach. Finally, similarly to the original approach of Gutzwiller, RMFT is devised only for $T = 0$.

Nonetheless, RMFT possess two generic properties of the MF approach, which turn to be important for the description of the cuprates. Namely, first, it allows for a natural and relatively simple description of various coexisting or competing symmetry-broken states, which are encountered in the cuprates. Stripe phases [73, 74, 75] or valence-bond solid [78] are good examples of the complex symmetry broken patterns that can be described within RMFT. Second, within this independent-particle picture, a Fermi surface (FS) appears in a natural manner, and the single-particle spectral properties can be easily addressed. Interestingly, the notion of the Fermi surface, being one of the most important concepts in solid state physics, is not limited to the non-interacting or weakly interacting systems. It is known from numerous photoemission experiments [87, 88, 89, 90, 91, 92, 93], that FS or FS-like structures are present in the cuprates, despite the presence of strong electron correlations.[3]

The question is how to modify RMFT in order to reproduce accurately the physical properties of the cuprates. Several attempts to improve the original formulation of RMFT have been made, cf. e.g. [63, 66, 71, 76]. Such modifications turned out to be important for a more realistic description of the RVB state.

1.2.1 Slave-boson theories of the t-J model

At this point we ought to mention another type of the MF approach, which is frequently used in the context of the t-J model, namely that based on the slave-boson formalism, i.e., the *slave-boson mean-field theory* (SBMFT). Historically, SB approach in general, and SBMFT in particular, where applied to the t-J model as early as in 1987 by Baskaran, Zou, and Anderson [94], by Baskaran, Anderson, Hsu and Zou [95], and Baskaran and Anderson [96], and later by Kotliar and Liu [97], and Suzumura, Hasegawa, and Fukuyama [98]. SMBFT techniques gained popularity, and those early papers were soon followed by many Authors.

Similarly to the RMFT, SBMFT provides a simple way for implementation of the RVB concept. Also, most versions of SBMFT lead to the predictions similar to those of the simplest realizations of RMFT approach. Moreover, the standard SBMFT approach is in fact equivalent to the properly treated corresponding version of RMFT, as discussed in Refs. [99, 100, 101] and also recently [102].[4] Apparently, SBMFT is a finite-temperature approach, in contrast to RMFT, which was devised to examine the ground state properties of the system. Yet, RMFT may be formally extended to $T > 0$, where, however, for various reasons both approaches

[3]We should rather say that the results of ARPES measurement are interpreted in terms of FS existence, e.g. by fitting the tight binding dispersion relation to the experimental data.

[4]Strictly speaking, this is the case for the RMFT [59] and the corresponding SBMFT [103] for the Hubbard model. In case of the t-J model, some differences between those two approaches appear, e.g. the kinetic energy is renormalized in a different way, i.e., $\sim x$ within SBMFT and $\sim 2x/(1+x)$ within the simplest version of RMFT. However, this technical detail is inessential. What is important here is that we can construct a MF model *completely equivalent* to that resulting from SBMFT without invoking sophisticated field-theoretic techniques.

are not expected to lead to physically meaningful results [100]. Therefore, SBMFT have no advantage over RMFT, and will not be discussed here.

Beyond the mean-field level, slave boson models provide a valuable tool for studying strongly correlated systems, as they form a basis for the effective gauge theories for the cuprates and heavy fermions [104, 105, 106]. However, this topic is outside the scope of the present Thesis.

1.2.2 RMFT versus VMC method

The results of RMFT are often compared with those of VMC approach. VMC method provides a valuable tool for studying strongly-correlated systems; applied to the cuprate superconductors it is known to yield a good semiquantitative description of the SC correlated state, cf. Refs. [45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 56, 57]. Within VMC one treats the double occupancy exclusion in an essentially exact way, and hence this method is often regarded as being superior to any MF treatment. However, properly constructed and solved RMFT may, at least in principle, lead to the results similar to those of the VMC. Moreover, RMFT has also some advantages over VMC approach. First, its results are not limited to small clusters. Second, it offers an analytic insight into the physical contents of the model and its relevance to the experiment.

1.2.3 Nonstandard character of RMFT approach

It is important to emphasize at this point, that the RMFT of the t-J model is not of the form of the standard Hartree-Fock (HF) MF approach. Therefore, a proper solution of RMFT, in particular of its more advanced versions, constitutes a nontrivial task. For the MF Hamiltonians of the HF form (cf. Section 4.9 for the precise definition of this term), minimization of the appropriate MF thermodynamic potential (the ground-state energy in particular) is equivalent to the approach based on the self-consistent equations (in the theory of superconductivity known under the name of Bogoliubov-de Gennes (BdG) equations). The latter express the basic requirement of the internal consistency of the mean-field model. The BCS theory [41, 42] is a good example of this equivalence. Also, for the HF MF Hamiltonians, the solutions of the MF model (i.e., the ground states of the MF Hamiltonian, corresponding to different patterns of symmetry breaking) provide us with the upper bounds on exact free energy (or the ground state energy in the $T \to 0$ limit). This is ensured by the Bogoliubov-Feynman inequality [107] and its generalizations [108] (cf. Section 4.10).

In general, neither of the last two statements is true for the RMFT approach. First, the unwary application of the variational method, i.e., direct minimization of the MF free or ground-state energy may lead to results that differ from those obtained by solving the self-consistent BdG equations. Moreover, by applying the Gutzwiller approximation, we may obtain values of the energy which are *lower* then the exact ground state energy of the original t-J model.

In such a situation, a non-variational treatment based solely on the BdG equations is frequently selected [63, 68, 73, 74, 75, 78]. However, this way of approach cannot be regarded as fully satisfactory.

2 Aim and a scope of the Thesis

A need for a consistent treatment of RMFT motivated us to develop a general method of solving mean-field (MF) models. Our approach is based on the *Maximum entropy principle* (MaxEnt) [109, 110, 111, 112], and may be regarded as a natural extension of the original formulation of this principle to the non-standard case of the MF approach. Construction of this formalism is the first principal aim of the present Thesis.[5]

The formal method of our approach is proposed in Part II, which is organized as follows. In Section 3 we comment on the origin, role and the nontrivial nature of MF methods in a general context. In Section 4 we present in detail the MaxEnt-based approach to MF models. In particular, in short Subsection 4.1, a notion of the MF model and MF Hamiltonian is formally introduced. Relation between the MaxEnt principle and a standard, non-MF statistical mechanics is reminded in Subsection 4.2, whereas the application of this principle in the context of MF statistical mechanics is discussed in Subsection 4.3. In Subsection 4.4 the optimal (equilibrium) values of mean-field variables and the correct form of the grand-canonical MF density operator are obtained. Subsection 4.5 is devoted to the construction of the Landau potential and to analysis of the non-equilibrium situation. Next, in Subsection 4.6 we discuss MF equilibrium thermodynamics. In Subsection 4.7 we analyze the role of chemical potential within MF description. In Subsection 4.8 we introduce a notion of equivalence class of the MF Hamiltonians. This and related concepts allow us, in particular, to reproduce formal results of other Authors within our approach. Subsection 4.9 is devoted to the important class of Hartree-Fock MF Hamiltonians, whereas in Subsection 4.10 we comment on relationship of the present MaxEnt-based variational principle to the variational principle based on the Bogoliubov-Feynman inequality. Subsection 4.11 contains additional remarks, which are intended to clarify certain aspects of the present formalism. Section 5 contains summary of Part II.

In Part III, the results of Part II are applied to the RMFT of the t-J model. We begin with the introduction of different forms of the t-J Hamiltonian and discussion of some of its general properties (Section 6). Next, in Section 7 we present various trial variational wave functions used as approximate ground states of the t-J Hamiltonian. It is shown, that a special class of such wave functions (so-called *correlated states*) leads in a natural manner to an effective, single-particle mean-field description in the form RMFT.

In Section 8, on the example of the simplest form of the t-J Hamiltonian, and by using different versions of RMFT approach, we compare first the results of the present variational approach with those of the non-variational treatment based on Bogoliubov-de Gennes self-consistent equations. The following MF states are analyzed: nonmagnetic, homogeneous superconducting state of a d-wave symmetry (dSC), (cf. e.g. Refs. [33, 39, 40, 58, 66, 61, 73, 74, 75, 82, 83], to mention just a few), staggered-flux non-superconducting solution (SF) (cf. e.g. [8, 33, 62, 75, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124]), and the so-called *Pomeranchuk instability* (PI) of the normal state, i.e., the spontaneous breakdown of the C_{4v} symmetry [33, 125, 126, 127, 128, 129, 130], cf. also Ref. [82].

On the example of those three states we show non-trivial differences between the results obtained by either different method, or different variant of RMFT. Next, the optimal form of RMFT is selected and applied within the framework of our method to study various forms of the t-J model (Section 9). It is also shown, that by making use of the RMFT based on the original formalism of Ref. [76], we can produce the results comparable to those of VMC and which are also in reasonable agreement with the experiment. This is the second principal aim of the present Thesis.

Some supplementary material is provided in Appendices and Supplements (Part IV). In

[5] Some of our early results may be found in Ref. [113]

Appendix A (Subsection 11.1) we show, that in the case of mean-field models, the method of Lagrange multipliers is indispensable for application of the MaxEnt principle. In Appendix B (Subsection 11.2) we provide the proof of equivalence of two alternative formulas for the second derivative of the thermodynamic grand potential. In Appendix C (Subsection 11.3) we explain the way in which different Landau potentials can be constructed for a given mean-field model. In Appendix D (Subsection 11.4) we present briefly some details of the formalism of Ref. [76]. Supplement A (Subsection 12.1) is devoted to the analysis of some aspects of the non-equilibrium situation, not discussed in Section 4.5. In particular, we discuss both thermodynamic and quantum fluctuations and the internal consistency of the mean-field approach. In Supplement B (Subsection 12.2), we analyze zero-temperature formulation of the MF approach. Finally, in Supplement C (Subsection 12.3), we illustrate the formalism developed in Section 4 on the example of the MF approach to Ising model.

As mentioned previously, RMFT description is not expected to be an equally legitimate approach within the entire x-T phase diagram of the cuprates. Consequently, we focus here only on the low-temperature situation and on the optimally-doped and overdoped regimes which are believed to exhibit a nonstandard, but essentially Fermi-liquid-type behavior. Therefore, we neglect any long-range magnetic order, in particular, simple Néel antiferromagnetic order. We also neglect any effects of the external magnetic field.

Within the model considered here, only a single CuO_2 layer is treated. In most cases (with the exception of the PI phase, analyzed in Section 8.3) we assume the presence of a discrete C_{4v} rotational symmetry. The superconducting order parameter $\Delta(k_x, k_y)$ is taken to be a singlet of $d_{x^2-y^2}$ symmetry, i.e., changes sign after a rotation of $\pi/2$ radians.[6] Therefore, we concentrate rather on generic features of the cuprates in the vicinity of the upper critical concentration x_c, although attempts to obtain material-specific results (by taking appropriate values of the model parameters) are also made. We analyze mainly the doping dependence of a gap magnitude and selected features of the quasiparticle spectrum in the superconducting state. The particular emphasis is put on x_c, which value is quite correctly predicted for the realistic values of the model parameters. This is the first such prediction within RMFT.

Although a consistent treatment of the RMFT of the t-J model was our original motivation, the formalism presented in Part II is of a general applicability, and may be used to treat wide class of the mean-field models. It has a number of advantages, not present in standard formulation of the MF theory. We hope than this method will be found useful in the condensed matter physics or even beyond the field. A work along these lines is being continued in our group.

Present Thesis is an extension of our earlier works [82, 83, 84, 113]. It contains (in a modified and refined form) main part of [113], large parts of [83] and essentially the whole material presented in [82]. Also, in Ref. [102] the formalism developed here has been used to show the equivalence of the mean-field approach resulting from the Gutzwiller approximation to the Hubbard model, with the corresponding slave-boson mean-field theory.

[6] Despite this particular form of $\Delta(k_x, k_y)$, the MF Hamiltonian, and hence the thermodynamic potentials, are still C_{4v}-symmetric.

Part II
Application of maximum entropy principle to mean-field models

3 Synopsis: qualitative aspects of mean-field theory

3.1 General remarks on mean-field approach

3.1.1 Introductory remarks

A rigorous treatment of even simple models of many interacting particles is usually too difficult. In such a situation, various approximate methods are to be developed.

We focus here on the so-called *mean-field* (MF) approach. Within the MF approach, the original, many-body Hamiltonian is replaced by its simplified MF counterpart, which becomes tractable. Instead of interacting with each other via full many-body potentials, the particles (or spins) are allowed to interact only with various 'mean fields' of semi-classical[7] character. Additionally, mean fields usually have an interpretation of average values of certain operators appearing in the MF Hamiltonian. Numerical values of such averages are not *a priori* known and are to be determined when solving a MF model.

From a historical perspective, methods of an essentially MF character were used first by van der Waals to derive equation of state for non-ideal gas (1873) [131, 132, 133], and next by Weiss to describe paramagnetic - ferromagnetic transition (1908) [132, 134, 135], both examples predate modern quantum mechanics (1925-1927). Probably the best-known example of the quantum MF approach is the Hartree-Fock (HF) approximation [136, 137], which has been used, in particular, in the Bardeen-Cooper-Schrieffer (BCS) theory of superconductivity (1957) [41]. However, mean-field methods (mainly in the form of the HF approximation) found numerous applications not only in the field of solid-state physics, but also in atomic [138], high-energy [139, 140, 141], and nuclear physics [137, 142], as well as in astrophysics (cf. e.g. [143]) and in quantum chemistry [144].

MF approach is still widely used, despite the fact, that other approximate methods exist, with the help of which we are able to treat interactions in a more accurate manner. This is partly due to the circumstance, that the MF approach is more direct and intuitive then most of the more sophisticated methods, which usually involve a greater emphasis on numerical analysis. Presence of explicit analytical formulas (e.g. for the ground state energy) frequently allows to make certain qualitative predictions, even before the MF model is completely solved. Also, MF methods are practically not limited by the system size. Therefore, MF approach is frequently the simplest available tool at hand, even if the proper solution of the MF model may also turn out to be a highly nontrivial task. However, apart from relative simplicity, there exist other, deeper and more subtle reasons determining the importance of the MF theory. This is discussed below, where we also invoke certain facts from both quantum and statistical mechanics.

3.1.2 Spontaneous symmetry breaking

By *spontaneous symmetry breaking* we understand a situation, when symmetry of the actual state of the system is lower then the symmetry of the Hamiltonian. In general, this means that the symmetry of the ground state is a subgroup of the total symmetry group of the Hamiltonian.

[7] An attempt to ascribe more precise meaning to this term will be made in what follows.

This concept plays a central role in many areas of modern physics [132, 133, 145, 146]. Apart from condensed matter physics, it is also widely used in the realm of high-energy physics and even in cosmology [147].

To describe spontaneous symmetry breaking in a quantitative manner, Landau introduced a concept of *order parameter* [148, 149]. By this term we understand any physical quantity, which has nonzero value in an ordered, i.e., *less* symmetric phase, and vanishes on the opposite side of the transition point or line. Magnetization (i.e., magnetic moment of a given volume of a specimen) may serve as a good example in the case of ferromagnetic-paramagnetic transition in the system of interacting spins.

3.1.3 Landau theory

A notion of an order parameter is fundamental for the theory of phase transitions, developed by Landau in the years 1936-1937 [148, 149]. Within this theory, generalized thermodynamic potential is introduced, which global minimum with respect to order parameter(s) corresponds to the equilibrium situation. In contrast to ordinary thermodynamic potentials encountered in standard thermodynamics (e.g. free energy or grand potential), generalized potentials of Landau theory[8] may depend on some thermodynamic variable and its conjugate variable (related by the Legendre transformation) at the same time. For example, in the case of magnetically ordered systems, generalized potential depends on both magnetization (an order parameter) and the external magnetic field. Only *after* the minimization is carried out (with respect to e.g. magnetization), the conjugate variables are no longer independent, e.g. the equilibrium value of magnetization is a function of magnetic field (and of other thermodynamic variables, such as the volume or temperature).

Landau theory, even if soon recognized to be quantitatively inaccurate (i.e., it predicts incorrect values of the critical exponents), had a great impact on theoretical physics [132]. It was later generalized by Ginzburg and Landau in order to provide a description of superconductivity [151]. Both Landau and Ginzburg - Landau theories in the original formulation have phenomenological character, which means that they make almost no assumptions about the underlying microscopic picture.

However, it may be expected, that there exists a close connection between Landau or Ginzburg - Landau theory and the microscopic MF models. Mean-field variables frequently play the role of order parameters, and the results of microscopic MF formulation, before the mean-fields optimization, are interpreted in terms of the Landau theory, cf. [146]. Also, development of Landau approach seems to have been (at least partly) motivated by the microscopic mean-field models existing at that time, e.g. the Bragg-Williams treatment (1934) of order-disorder transition in binary alloys [152]. On the other hand, following Gorkov (cf. e.g. [44]), one may start from the microscopic MF model and derive the corresponding Ginzburg-Landau functional by applying Green's function technique within the BCS theory.

In the present Thesis a natural connection between phenomenological description in the spirit of Landau and Ginzburg, and the microscopic MF models will be established from a different perspective.

3.1.4 Mean-field approach as a semi-classical description

Apparently, there exists some relationship between spontaneous symmetry breaking and emergence of the classical world from the laws of quantum mechanics.[9]

[8] Here by 'Landau theory' we always understand the theory of phase transitions, and not the theory of Fermi liquids [150]. For the latter, the full name 'Landau theory of Fermi liquids' is always used.

[9] Highly non-trivial relationship between quantum and classical physics is still not resolved. There are many attempts to solve this problem, e.g. by invoking environmental decoherence (cf. [153] and References therein).

Obviously, symmetries of classical, macroscopic objects are very different from those which are present on the quantum level. Following Ref. [157], consider an example of crystalline state. Microscopic quantum Hamiltonian describing a collection of atoms (or ions) and electrons exhibits full translational invariance, yet in crystals the full translational symmetry is broken, and the atoms can form a regular structure.[10]

Apart from appearance of the crystalline state, for which symmetry-breaking is evident, essentially the same situation appears in the cases of magnetic ordering or appearance of superconductivity. For example, let us consider Neel state, characterized by a static, long-range antiferromagnetic order. This state is obviously *not* an exact eigenstate of the Heisenberg Hamiltonian (cf. (6.127) in Section 6.135 and (12.282) in Section 12.3), regarded as a minimal microscopic model of real antiferromagnetically ordered materials. As another example, we may consider a ferromagnet, which may also be described by a ground state of the Heisenberg Hamiltonian, with the opposite (negative) sign of the exchange integral. However, the point is, that in the absence of an external magnetic field, this ground state is highly (strictly speaking, infinitely) degenerate. By selecting a direction of the spontaneous magnetic moment we break the $SU(2)$ symmetry of the quantum model.

Using the above examples, we conclude, that the standard (i.e., without the concept of the symmetry breaking) quantum-mechanical treatment may be inconvenient, or even insufficient to describe various symmetry-broken states of matter, frequently encountered in condensed matter physics. Such states, whose existence is experimentally evident, do not correspond to eigenstates (or at least to unique eigenstates) selected out of the states of those quantum Hamiltonians, which are regarded as defining correct and essentially complete microscopic models of the systems in question. Now, it should become more clear, why the symmetry-broken states are sometimes termed '(semi-) classical' states or 'classical condensates' [157, 158].[11] They exhibit peculiar properties, in particular long-range order and 'rigidity', i.e., robustness with respect to external perturbations.

Interestingly, 'classical condensates' can be modeled using eigenstates of the appropriate MF Hamiltonians. Existence of long-range order(s) is build into such description in a natural manner, and existence of the 'classical domain' is *a priori* assumed. This means, that when solving a MF model, we determine the actual optimal values of mean-fields, which may indeed differ from zero. By doing so, we usually break some of the unitary symmetries originally present in the microscopic MF Hamiltonian.

Non-zero values of mean-fields may imply that there exists a finite gap in the spectrum of the MF Hamiltonian. The presence of the gap, in turn, explains the rigidity; the system remains in the ground state despite the external perturbations, as long as the latter are weak enough (i.e., characterized by the energy scale smaller than a gap).[12]

Due to non-zero value of the gap, a difference between pure, single-determinant ground state of the MF Hamiltonian, and a mixed thermal state is insignificant at low temperatures. However, mean-field models could be also used to describe the symmetry-broken states at non-zero temperature, and then obviously both the ground state, and the excited states of the MF Hamiltonian are required. In the present Thesis, we propose an approach to MF models valid for arbitrary $T > 0$.

However, the latter point of view has been critically examined [154, 155, 156]. This fascinating topic is outside the scope of the present work.

[10] Obviously, the existence of crystals, or any other macroscopic objects localized in space, also breaks translational symmetry.

[11] In Ref. [157] a precise distinction between classical and semi-classical states is made, but we do not follow this terminology strictly.

[12] It should be noted that the gap existence is not the necessary, but rather the sufficient condition for the rigidity of the broken symmetry state. For example, there exist zero-gap superconductors, in which the phase rigidity of the macroscopic wave function is the principal factor.

3.1.5 MF formalism as a result of saddle-point approximation

There exits yet another aspect of the classical character of the MF approach. Namely, it is well-known, that the quantum-mechanical description of a single, spinless particle may be formulated in terms of path integrals [159]. Conversely, given a quantum-mechanical propagator, we may distinguish a stationary path, corresponding to the classical trajectory (i.e., one making the classical action stationary). Similarly, making use of the coherent states of spin, Bose or Fermi operators, one may express the partition function of a many-body system as a path integral [160]. Following Fradkin [161], let us consider Hubbard model as an example. One may apply Hubbard-Stratonovich (HS) transformation when determining the relevant partition function. This step leads to an equivalent problem, expressed in terms of both fermionic and auxiliary bosonic fields. Importantly, the partition function expressed in this way is quadratic in fermionic degrees of freedom, therefore the latter can be integrated out, and one obtains an effective (Euclidean) action in terms of auxiliary Bose fields introduced by HS transformation. It turns out, that a saddle-point approximation applied to such an effective action is equivalent to the Hartree-Fock mean-field approximation [158]. In analogy to the single-particle case, the path, singled out by means of the saddle-point approximation, is said to correspond to a 'classical' situation. This is less obvious in the case of many-body system, than for a single particle, but it still seems to be justified to call the Hartree-Fock approach a 'semi-classical theory' [158].

Following Refs. [157, 158], we want to point out here, that it is not the weakness of the interaction, which justifies MF approach, but rather the existence of a non-zero value of certain order parameter(s) and a subsequent ability to include the fluctuations around the mean-field solution to obtain a complete description. In other words, description in terms of the MF states may be regarded as more than just an approximation to the proper ground- or equilibrium-state description of some many-body Hamiltonian in the weak-coupling regime, even if this is the role the MF states often play.

3.1.6 Mean-field approach as a description based on restricted class of quantum observables

As pointed out in Refs. [162, 163, 164], mean-field theory may also be regarded as an attempt to describe a physical system by using only quantum operators, which belong to some restricted class. In the case of fermionic system encountered in the condensed-matter physics, this usually mean that we use operators which are bilinear in creation or annihilation operators, i.e., our MF Hamiltonians are of single-particle nature. The latter choice is privileged in connection with the application of Wick's theorem [136, 165], but other classes of operators may be preferable e.g. for the mean-field models of bosonic systems (e.g. for bosons in optical lattices [166]) or the mean-fields models used in nuclear physics [162, 164].

3.1.7 Statistical mechanics and spontaneous symmetry breaking

It is well known, that for finite systems, standard statistical mechanics does not predict neither temperature-driven phase transitions, nor the spontaneous symmetry-breaking. Indeed, at the phase transition point the thermodynamic potentials must be non-analytic functions of the inverse temperature $\beta \equiv 1/k_B T$. On the other hand, partition function of the finite system is a sum of *finite* number of terms of the form $\exp(-\beta E_i)$ or $\exp(-\beta(E_i - \mu N_i))$. Each such term, as well as their finite sum is an analytic function of the inverse temperature, therefore we can never obtain true, 'sharp' phase transition [132]. Also, in the absence of an external, symmetry-breaking field, all the micro-states related by the symmetry transformations, with respect to which the Hamiltonian is invariant, have the same energies and enter the partition

function with the same weight. As a consequence, order parameters, which are averages of certain microscopic quantities, cannot retain non-zero values. Ising model [167] in the absence of the external magnetic field may serve as an example. Because each micro-state has its spin-reversed partner with exactly the same energy, total magnetization is equal zero.

Therefore, within a standard statistical mechanics, phase transitions and the symmetry-broken states of matter seem to be intrinsically related to the presence of large, or strictly speaking, infinite number of microscopic constituents of the system, i.e., to the thermodynamic limit [132].

3.1.8 Method of quasi-averages

Symmetry-broken quantum states may be obtained by means of the method of quasi-averages proposed by N. N. Bogoliubov [168, 169]. Quasi-averages are defined in a thermodynamic limit, and in the presence of an additional, external symmetry-breaking field. Eventually, this field is turned off *after* the thermodynamic limit is taken. Importantly, the order of those two operations cannot be interchanged [145, 168, 169, 170, 157].

However, for many exact (non-MF) models of particular interest, it may be rigorously shown, that depending on the spatial dimensionality D_S and temperature T, the method of quasi-averages does not yield the symmetry-broken states with a true long-range order. Notable examples are: a lack of the long range superfluid (superconducting) order in Bose (Fermi) liquids for $D_S = 1$ and $D_S = 2$ at $T > 0$ [171], or a lack of antiferro- and ferromagnetic ordering in the Heisenberg model in $D_S = 1$ (at $T \geq 0$) and $D_S = 2$ (at $T > 0$) [172]. Moreover, various symmetries of the superconducting order parameter are excluded in the two-dimensional Hubbard model [173, 174, 175]. Also, as pointed out in [176], there is even no superconductivity of a $d_{x^2-y^2}$-wave symmetry in the two-dimensional t-J model, commonly regarded as a correct minimal model of the high-temperature cuprate superconductors.

Interestingly, in each of the above mentioned cases, the corresponding MF approximations yield the symmetry-broken solutions easily, and quite insensitively to the system size, dimensionality or temperature. In general, MF approach overestimates range of ordered phases, e.g. it yields critical temperatures which are far too high. On the other hand, MF methods allow us to describe symmetry breaking in real systems using simple, low-dimensional models. As discussed above, this usually would not be case for an exact treatment of full many-body problem, even if such treatment was technically feasible.

3.1.9 'More is different'

We may look at the previous discussion, concerning the existence of the 'classical condensates' and insufficiency of quantum mechanics to describe such states, from even more general perspective. Namely, following Anderson [145], let us note that it may be technically or conceptually impossible to predict the collective behavior of complex systems, even if we have a complete knowledge about the interactions between their microscopic constituents. Existence of a non-zero dipole moment of certain molecules, like ammonia NH_3 and its heavier analogs, (e.g. phosphine, PH_3) is a striking example given by Anderson. However, this situation is encountered not only in chemistry or in solid state physics. Even apparently more fundamental theories have some phenomenological ingredient build in [145, 157]. For example, in quantum chromodynamics (QCD), a kind of a MF approach is used to explain the origin of mass of nucleons ('chiral condensate') [139].

3.2 How to solve mean-field models?

Before we may answer this question, first let us define what we mean by 'solution of the MF model'. Namely, MF model is solved once the optimal values of mean-fields are determined and the explicit form of the MF density operator is known. Obviously, these two goals are closely related. MF density operator is required in order to compute expectation value of any operator, which may be relevant to the problem at hand. At the same time, MF density operator depends functionally on the mean-field variables. Note, usually the diagonalization of the MF Hamiltonian is rather straightforward. However, it may be quite problematic, what do we understand by 'the optimal values of mean-fields'.

3.2.1 Variational principle based on Bogoliubov-Feynman inequality and its generalizations

At $T > 0$, MF density operators are frequently used as trial variational states within the variational principle based on the Bogoliubov-Feynman inequality [107] and its generalizations [108], cf. also Subsection 4.10. Using the Bogoliubov-Feynman inequality, we obtain an upper bound for the grand potential Ω_e or free energy F_e of the system described by some non-MF ('exact') Hamiltonian $\hat{H}_e$.[13] In other words, from such point of view, MF Hamiltonians and MF states play only an auxiliary role. However, if the mean-field variables are treated as variational parameters, their optimal values obtained from Bogoliubov-Feynman principle are in general *not* equal to the averages of the corresponding operators, contrary to basic definitions of mean-fields (we comment more on this point in Subsection 4.10).

One may argue, that what should mainly concern us is the optimal upper bound for the corresponding thermodynamic potential. Therefore, the internal self-consistency of the MF model would be of secondary importance and may be ignored. However, in our opinion, this point of view is unacceptable.

On the other hand, even if the value of free (or the ground state) energy of the MF model is close to the exact one, it is not guaranteed at all, that the original many-body ('exact') model and its MF counterpart are similar with respect to any other property. In such a situation, one may try to use a dedicated variational principle suited to the optimization of each quantity of interest [177]. However, in such a case we simultaneously deal with several *different* variational principles; one of them is variational principle for the free energy, based on the Bogoliubov-Feynman inequality. Therefore, in the context of the MF theory, the formalism of Ref. [177] leads to situation which is qualitatively similar (though technically more complicated) to that resulting from the application of Bogoliubov-Feynman principle. This route thus not seem to be the preferable way of solving MF models.

3.2.2 Mean-field description involving only a mean-field Hamiltonian

In the previous paragraphs we have discussed some non-trivial features of the MF formalism. Namely, we have pointed out that MF models are able to provide a description of the symmetry-broken phases, which is frequently not the case for the corresponding non-MF models. In other words, MF models should be considered as being more than just crude approximations to intractable non-MF models. Therefore, it seems legitimate to consider a situation, when the system is described *entirely* in terms of the MF Hamiltonian. Namely, although some 'exact' (non-MF) Hamiltonian $\hat{H}_e$ may be initially postulated for the problem, (and even being used to construct the MF Hamiltonian in question), it is eventually disregarded and should not be required to solve the MF model. In particular, neither the Bogoliubov-Feynman inequality nor

[13] In $T \to 0$ limit, the Bogoliubov-Feynman principle reduces to the variational principle of quantum mechanics, and we obtain upper bound for the ground state energy.

its generalizations ($T > 0$), nor the minimization of the expectation value of $\hat{H}_e$ computed with respect to the mean-field quantum state ($T = 0$), should be invoked. Our motivation for solving the mean-field models 'on their own' partly comes from the Landau theory. However, this way of dealing with the MF theory seems to be logical and consistent in its own right.

We are aware, that such point of view on the MF approach may be not convincing to all or even regarded as questionable. Namely, one may argue, that without a non-MF Hamiltonian $\hat{H}_e$ we have neither real basis nor a reference point for the analysis of the physical system. For example, depending on the MF Hamiltonian as well as the way we solve a MF model, the resulting value of the MF free (ground-state) energy may be even *lower* then the corresponding energy of this $\hat{H}_e$, which is considered as realistic and correct microscopic model of the system in question. Nonetheless, apart from the arguments given previously, please note that approaches based *de facto* entirely on the MF Hamiltonian are still widely applied to study various systems encountered in condensed matter physics. In particular, the MF models based on the Gutzwiller approximation (GA) [20, 59] are frequently used to provide an approximate treatment of the models of strongly correlated fermions, i.e., various forms of the Hubbard model, the related t-J model within the RMFT picture [33, 58], or of the Anderson model. Similar situation is present in case of the saddle-point (MF) approximation to various forms of the slave-boson formalism [33, 103], which also leads to an effective, single-particle picture. In this Thesis, most of the just raised points are subject of detailed analysis on example of a concrete model.

3.2.3 Approach based on Bogoliubov-de Gennes equations

RMFT of the t-J model may serve as an example of the mean-field theory which is not of the standard Hartree-Fock form (as defined in Subsection 4.9). In result, the Bogoliubov-Feynman variational principle is not equivalent to the minimization of the MF grand-potential potential with respect to the mean-field variables (cf. discussion in Subsection 4.10). Moreover, for more advanced versions of RMFT, both variational procedures spoil the self-consistency of the model. In that situation, RMFT is frequently solved by invoking the basic self-consistency conditions, leading to the *self-consistent* equations (in the case of the mean-field treatment of superconductivity termed also *Bogoliubov-de Gennes* equations) for the values of mean-fields (cf. Eqs. (4.8) of Subsection 4.3), and no variational procedure is used. However, this non-variational way of solving MF models, although internally consistent, suffers from many drawbacks. In fact, a need for a more satisfactory treatment of RMFT was the original motivation for the construction of the general formalism presented in this Thesis.

3.2.4 Maximum entropy principle

If we decide to base our description entirely on the MF Hamiltonian, its eigenstates represent all available micro-states of the system. In other words, we try to 'mimic' an exact description of the system, originally based on some non-MF many-body Hamiltonian $\hat{H}_e$, by means of the properly chosen MF Hamiltonian. In order to solve MF model at $T > 0$, we need to determine probability distribution for the microstates of the system. We are concerned mainly with the equilibrium properties. Accordingly, we construct appropriate canonical or grand canonical ensemble for the MF situation.

We base our approach on the *maximum entropy* (MaxEnt) *principle* [109, 111, 110]. In general, the MaxEnt principle is one of the cornerstones of Bayesian mathematical statistics. It allows to construct the optimal probability distribution on the basis of incomplete information [109, 111, 110]. In particular, it may also be regarded as the basis of statistical physics, and the latter is treated then as a special case of statistical inference. From this point of view, canonical density operators of equilibrium statistical mechanics are derived using MaxEnt

principle [109, 112, 178, 179]. In fact, this route is not only the most convenient, but also the most general one. It could be relatively easily generalized to treat various non-standard situations, e.g. the non-equilibrium statistical mechanics [178]. We argue that the MaxEnt principle should be also used in the case of a mean-field description of the system. However, in order to carry out the whole procedure, a standard formulation of the MaxEnt principle must be modified.

It turns out, that such approach to the MF models has a number of attractive features. First, it allows for a natural interpretation of a given microscopic MF model in terms of the Landau theory of phase transitions. Namely, for each MF Hamiltonian, a corresponding Landau potential may be constructed explicitly. Consequently, not only the MF equilibrium thermodynamics is defined in a consistent manner, but the non-equilibrium situation can also be studied. Although we consider only the stationary, i.e., time-independent situation, such analysis may help to reveal the internal limitations of the present MF approach.

We may hope that our description of the equilibrium situation is, by its construction, optimal from the point of view of mathematical statistics. Any other probability distribution one may find for a given MF model would be biased, i.e., would contain information which we in fact do not possess, or some available information would be ignored.

Let us note here, that an approach to MF theory based on the maximum entropy inference, has been also addressed in the series of papers [162, 163, 164]. Existence of those References apparently has not been noticed by a condensed matter community. We too became aware of it only after completing the main parts of the present work. Formalism of those References differs from ours, nonetheless, both approaches share also some common features. We provide some discussion of that point in Section 4.8 and in the Summary of Part II.

3.2.5 Optimal effective mean-field picture

It is important to note here, that although the method we propose has been initially devised to provide the unbiased solution of an *a priori* given MF model, it may be also used to construct the optimal (from the point of view of MaxEnt inference) approximation to a given non-MF many-body Hamiltonian $\hat{H}_e$. In this manner the rigorous upper bound on the exact free energy, or (in the $T \to 0$ limit) the ground state energy of $\hat{H}_e$ may be obtained in principle. Also, our approach may be combined with the one based on the Bogoliubov-Feynman inequality, in such a way, that the self-consistency of the MF approach is preserved.

3.2.6 Zero temperature situation

Finite-temperature MaxEnt-based formalism constructed here offers also a convenient starting point for the analysis of $T = 0$ situation. The latter may be recovered as the $T \to 0$ limit of the $T > 0$ case. Such an approach omits many of the difficulties of pure zero-temperature treatment, as discussed in detail in Sec. 12.2 (Supplement B).

Several examples of zero-temperature, self-consistent variational MF approach can be found in the literature, cf. e.g. [66, 69, 70, 71, 85]. In each of those References, the Authors make use of the variational principle of quantum mechanics (minimization of the expectation value of an appropriate Hamiltonian) for the MF case, and MaxEnt principle is not invoked. Moreover, the results of Refs. [66, 69, 70, 71, 85] may be obtained within our formalism in the $T \to 0$ limit.

3.3 Summary of synopsis

In summary, the general character and the deductive nature of the modified MaxEnt approach justifies, in our opinion, its detailed exposition in this Thesis before concrete models are tackled

explicitly. Most of the results presented in the next Section will be later utilized in Part III. Topics less related to the contents of Part III are presented in the Appendices and Supplements (Part IV).

4 Formalism and method

4.1 Mean-field Hamiltonian

Consider a *mean-field* (MF) Hamiltonian,

$$\hat{H}(A_1, A_2, \ldots, A_M) \equiv \hat{H}(\vec{A}), \tag{4.1}$$

which depends on M *mean fields* $A_1, A_2, \ldots, A_M$, $(A_1, A_2, \ldots, A_M) = \vec{A}$.[14] We assume that the explicit form of $\hat{H}(\vec{A})$ is *a priori* known. Each A_s has an interpretation of expectation value of the corresponding operator, i.e., $A_s \equiv \langle \hat{A}_s \rangle$. The averages $\langle \ldots \rangle$ are defined in the standard manner, i.e.,

$$\langle \hat{A} \rangle = \mathrm{Tr}[\hat{A}\hat{\rho}] = \mathrm{Tr}[\hat{\rho}\hat{A}], \tag{4.2}$$

where $\hat{\rho}$ is the MF-density operator, the explicit form of which is not specified as yet, and has to be subsequently determined. We assume that $\hat{A}_s$ operators are time-independent in the Schrödinger picture, i.e., $\partial \hat{A}_s / \partial t = 0$. Also, without loss of generality, we may assume that $A_s \in \mathbb{R}$ (this can always be achieved by an appropriate choice of $\hat{A}_s$ operators). Those $\vec{A}$, for which $\hat{H}(\vec{A})$ is well-defined will be called a *domain* of the MF model and denoted by $\mathcal{D}_{\mathcal{A}}$.

MF Hamiltonian, together with the lattice of spatial dimension D_S and particular geometry define a MF model. In order to avoid inessential mathematical difficulties related to the underlying algebraic structure, we assume that the dimension D_H of the Hilbert space, in which $\hat{H}(\vec{A})$ is defined, is finite, $D_H < \infty$. Here such restriction is acceptable, as we are mainly interested in the fermion- or spin models defined on finite (albeit arbitrary large) lattices.[15]

We also assume that $\hat{H}(\vec{A})$ does not depend explicitly either on time, or on purely statistical quantities like temperature or chemical potential, unless explicitly stated otherwise.[16] At this point we impose no further restrictions on $\hat{H}(\vec{A})$, and we are not concerned with the problem of derivation or physical justification of its particular form. However, it turns out that the method developed here may be useful also for that purpose, see Subsection 4.8.

Our aim here is to describe a physical system only by making use of the mean field Hamiltonian $\hat{H}(\vec{A})$. We assume that the system is in contact with a heat bath and possibly with some particle reservoir.[17] Therefore, we will need grand-canonical (GC) ensemble for the case of MF description of the system. Obviously, GC ensemble is preferred for lattice fermion and boson models. This feature is even more pronounced in the case of MF models, than for the standard statistical mechanics, because the average particle number $\langle \hat{N} \rangle$ is usually one of the relevant mean-fields. To avoid confusion, $\langle \hat{N} \rangle$ will be denoted by A_1, and not by N; the latter symbol will be used only for the equilibrium value of the average particle number, appearing in thermodynamic relations.

Other cases, i.e., micro-canonical and canonical ensembles may be treated analogously (cf. Sec. 12.3, where the MF model of the spin system is analyzed within canonical ensemble). Moreover, classical lattice models (e.g. Ising model) may be studied using a formally quantum-mechanical approach developed here, for the special case of Hamiltonians constructed from mutually commuting observables.

[14] We denote mean-fields as a components of a vector $\vec{A}$ mainly in order to have a compact notation. Nontrivial vector character of $\vec{A}$ will be discussed at the end of Subsection 4.8.

[15] An extension to the case of infinite, but countable Hilbert space is possible in principle. This would allow us to treat also the lattice-boson systems. In contrast, the case of uncountable basis cannot be treated within the present formulation of the method, as then the maximum entropy principle itself must be modified.

[16] It is sometimes convenient to relax the latter requirement and allow for an explicit T-dependence of the MF Hamiltonian.

[17] For simplicity, we assume that only one kind of (fermionic) particles is present. General case may be treated analogously.

4.2 MaxEnt principle and statistical mechanics

Before turning to the MF situation, let us remind briefly one particular way in which density operators of canonical statistical ensembles may be derived within a standard (non-MF) statistical mechanics. This route seems to be the most general as well as the most convenient one; it also turns out to be particularly convenient in the case of the MF models.

As pointed out by Jaynes [109, 110, 111], canonical density operators may be derived by invoking *maximum entropy principle* (MaxEnt) (see also [112, 178]). As an example, let us consider grand-canonical (GC) ensemble. If $\hat{H}_e$ is the Hamiltonian of the system, then the following functional

$$\mathcal{S}_e = \mathrm{Tr}\big(- \hat{\rho}_e \ln \hat{\rho}_e - \beta \hat{\rho}_e \hat{H}_e + \mu\beta \hat{\rho}_e \hat{N} - \omega \hat{\rho}_e \big), \tag{4.3}$$

has the maximum value for true GC density operator $\hat{\rho}_e$. The first term on the r.h.s. of (4.3) is the von Neumann entropy

$$S_{vN}(\hat{\rho}) = -\mathrm{Tr}(\hat{\rho} \ln \hat{\rho}), \tag{4.4}$$

whereas the other terms are related to expectation value of the Hamiltonian ($\langle \hat{H}_e \rangle_e \equiv \mathrm{Tr}[\hat{\rho}_e \hat{H}_e]$) and particle number operator ($\langle \hat{N} \rangle_e \equiv \mathrm{Tr}[\hat{\rho}_e \hat{N}] \equiv N$). The last term ensures correct normalization ($\mathrm{Tr}[\hat{\rho}_e] = 1$). The inverse temperature $\beta = 1/k_B T$, and the chemical potential μ play role of the Lagrange multipliers, enforcing the constraints for $\langle \hat{H}_e \rangle_e$ and $\langle \hat{N} \rangle_e$, respectively.[18]

In equilibrium, density operator $\hat{\rho}_e$ should be a constant of motion (we assume also that $\hat{H}_e$ does not depend explicitly on time, i.e., $\partial \hat{H}_e / \partial t = 0$). Therefore, the quantum Liouville equation (von Neumann equation), i.e.,

$$i\hbar \frac{\partial}{\partial t} \hat{\rho}_e(t) = [\hat{H}_e, \hat{\rho}_e(t)] \tag{4.5}$$

implies that $\hat{\rho}_e$ should commute with $\hat{H}_e$. As a consequence, and because also the condition $[\hat{H}_e, \mu\hat{N}] = 0$ usually holds for non-MF Hamiltonians, $\hat{\rho}_e$ and $\hat{H}_e$ have common eigenbasis. Next, we vary $\mathcal{S}_e$ (4.3) with respect to the diagonal elements of density operator, i.e., probabilities, and obtain the necessary conditions for the extremum of $\mathcal{S}_e$ (4.3) with respect to $\hat{\rho}_e$. This yields GC density operator of the form

$$\hat{\rho}_e = Z^{-1} e^{-\beta(\hat{H}_e - \mu\hat{N})}, \qquad Z = \mathrm{Tr}[e^{-\beta(\hat{H}_e - \mu\hat{N})}]. \tag{4.6}$$

It turns out that $\hat{\rho}_e$ (4.6) indeed corresponds to the maximum of (4.3).

Now, let us consider any operator $\hat{A}$ which does not depend on time explicitly (in the Schrödinger picture), i.e., $\partial \hat{A}(t)/\partial t = 0$. It could be easily verified, that then the average value of $\hat{A}$ in the mixed state (4.6) is time-independent. Indeed,

$$\begin{aligned} \langle \hat{A} \rangle_t &= \sum_i \langle i(t) | \hat{\rho}_e \hat{A} | i(t) \rangle = \sum_i \langle i(0) | e^{i\hat{H}_e t/\hbar} \hat{\rho}_e \hat{A} e^{-i\hat{H}_e t/\hbar} | i(0) \rangle \\ &= \sum_i \langle i(0) | e^{-i\hat{H}_e t/\hbar} e^{i\hat{H}_e t/\hbar} \hat{\rho}_e \hat{A} | i(0) \rangle = \sum_i \langle i(0) | \hat{\rho}_e \hat{A} | i(0) \rangle = \langle \hat{A} \rangle_{t=0}. \end{aligned} \tag{4.7}$$

[18] When applying the MaxEnt principle in general inference problems of Bayesian statistics, we usually encounter the following situation. A prior information about averages of certain functions (or operators in the quantum case) is available, and taken into account by imposing appropriate constraints. The Lagrange multipliers are determined from the requirement, that the resulting probability distribution is consistent with the knowledge we *a priori* have. This is exactly the case of $\langle \hat{N} \rangle = N$ and μ. However, for $\langle \hat{H}_e \rangle_e \equiv U$ and β a situation is different due to the physical interpretation of β. Namely, for an open system in thermal equilibrium, it is usually the temperature, and not the average energy which is experimentally accessible and therefore *a priori* known. However, those two ways of approach (i.e., either by fixing U or β) are equivalent. More detailed discussion of this point may be found in [109].

Above we have used only the fact, that $\hat{\rho}_e$ commutes with $\hat{H}_e$ and is time-independent. This holds, in particular, when $\hat{\rho}_e$ is an operator-valued function of the time-independent Hamiltonian, as in the present case.

4.3 Maximum entropy principle in the context of mean-field theory

4.3.1 Mean-field density operator and self-consistency conditions

Now we come back to the MF approach. It may be expected, that a nontrivial $\vec{A}$-dependence of the MF Hamiltonian $\hat{H}(\vec{A})$ (4.1) is also inherited by MF density operator $\hat{\rho} = \hat{\rho}(\vec{A})$. Regardless the detailed form of the latter, using Eq. (4.2) we should be able to determine, in particular, value of each of the mean-fields $A_1, A_2, \ldots, A_M$ appearing in $\hat{\rho}$, by solving the following equations

$$\langle \hat{A}_s \rangle = A_s = \mathrm{Tr}[\hat{A}_s \hat{\rho}(A_1, A_2, \ldots, A_M)], \qquad s = 1, 2, \ldots, M. \tag{4.8}$$

Eqs. (4.8) are termed *self-consistency equations* or *Bogoliubov - de Gennes* (BdG) equations. They guarantee that mean-fields are indeed average values of the corresponding operators. Obviously, the basic requirement expressed by Eqs. (4.8) should not be ignored, even if this is sometimes the case.

4.3.2 Incomplete treatment

One may expect, that MF grand-canonical (GC) density operator should be of the form (4.6), but with $\hat{H}_e$ replaced by $\hat{H}(\vec{A})$. However, this may be the case not. Namely, below we show that a construction leading to such form of GC $\hat{\rho}$ is incorrect in general.[19]

In the present situation we proceed analogously to the standard, non-MF case, i.e., we try to make use of the MaxEnt principle. In order to do that, we rewrite[20] (4.3) with $\hat{\rho}_e$ replaced by MF density operator $\hat{\rho}$, i.e.,

$$\mathcal{S} = \mathrm{Tr}\left(-\hat{\rho} \ln \hat{\rho} - \beta \hat{\rho} \hat{H}(\vec{A}) + \nu \hat{\rho} \hat{N} - \omega \Big(\hat{\rho} - \frac{1}{D_H} \Big) \right), \tag{4.9}$$

At first, let us forget for a moment about the interpretation of mean-fields as given by (4.8) and treat each A_s as some additional, non-variational parameter, which is independent from probabilities. Additionally, we assume that for all $\vec{A} \in \mathcal{D}_{\mathcal{A}}$, $\hat{\rho}(\vec{A})$ is time-independent. Invoking Eq. (4.5) for the present situation, it follows that $\hat{\rho}(\vec{A})$ commutes with $\hat{K}(\vec{A}) \equiv \hat{H}(\vec{A}) - \mu \hat{N}$. Then (4.9) may be rewritten in the common eigenbasis of $\hat{\rho}(\vec{A})$ and $\hat{K}(\vec{A})$, denoted by $\{|\tilde{i}\rangle\}_{i=1}^{D}$. This yields

$$\mathcal{S} = \sum_{i=1}^{D_H} \Big\{ - p_i \ln p_i - \beta p_i \Big([\hat{H}(\vec{A})]_{ii} - \mu [\hat{N}]_{ii} \Big) - \omega (p_i - \frac{1}{D_H}) \Big\}. \tag{4.10}$$

For any operator $\hat{O}$, by $[\hat{O}]_{ii}$ we denote $\langle \tilde{i} | \hat{O} | \tilde{i} \rangle$. Next, we maximize (4.10) with respect to each p_i, $i = 1, \ldots, D$, with the condition $\sum_j p_j = 1$ imposed (through $\partial_\omega \mathcal{S} = 0$). In effect, we indeed obtain a density operator of the form (4.6) but with $\hat{H}(\vec{A})$ replacing $\hat{H}_e$, i.e.,

$$\hat{\rho}(\vec{A}) = \tilde{Z}^{-1} \exp\big(- \beta (\hat{H}(\vec{A}) - \mu \hat{N}) \big), \qquad \tilde{Z} = \mathrm{Tr}[\exp\big(- \beta (\hat{H} - \mu \hat{N}) \big)]. \tag{4.11}$$

Eventually, after (4.11) is obtained, we may come back to the interpretation of each mean-field $A_1, A_2, \ldots, A_M$ as an average value of the corresponding operator $\hat{A}_s$. Then, those

[19] With the exception of MF Hamiltonians of the Hartree-Fock approximation, cf. Subsection 4.9

[20] We have changed the form of the normalization constraint as compared to (4.3).

parameters have to be determined self-consistently by using Eqs. (4.8). Thus, we obtain $A_1 = A_{1sc}, \ldots, A_M = A_{Msc}$, i.e., $\vec{A} = \vec{A}_{sc}$. However, the point is that (4.11) with $\vec{A} = \vec{A}_{sc}$, i.e., $\hat{\rho}(\vec{A}_{sc})$ is not true grand-canonical MF density operator for $\hat{H}(\vec{A}_{sc})$, even it has an apparent GC form. Obviously, this is because (4.11) has been obtained by neglecting the implicit dependence of $\vec{A}$ (hence, also $\hat{H}(\vec{A})$) on $\hat{\rho}$ due to (4.8). This step is often (implicitly) made (cf. e.g. [68, 73, 74, 75, 78] and [63] for the corresponding approach at $T = 0$), making the whole approach not fully consistent.

4.3.3 An attempt to eliminate mean-fields

In order to fix the incompleteness of the approach discussed above, we may try to express the expectation values A_s through the corresponding matrix elements of $\hat{A}_s$ operators and the probabilities $p_1, p_2, \ldots p_{D_H}$, using (4.8), i.e.,

$$\langle \hat{A}_s \rangle = A_s = \sum_i p_i \langle i|\hat{A}_s|i\rangle. \tag{4.12}$$

The probabilities become then the only independent variables present in the problem (modulo the normalization constraint). However, the mean-fields can be explicitly eliminated in favor of the probabilities only if the eigenbasis of $\hat{\rho}$, $\{|i\rangle\}_{i=1}^{D_H}$, is $\vec{A}$-independent. Even if this is the case, such route is very inconvenient. $\mathcal{S}$ given by Eq. (4.9) is then no longer of the canonical (Gibbs) form due to the term $\sim \langle \hat{H}(\vec{A})\rangle$, which is nonlinear in p_i. Consequently, the necessary conditions for $\mathcal{S}$ to have an extremum lead then to a set of nonlinear equations, which usually cannot be solved analytically. Apart from facing a problem of the numerical solution of large system of the nonlinear algebraic equations, we have no analytical formulas at our disposal, and thus, no deeper insight into the general properties of such constructed MF formalism. Moreover, in a general situation, the $\vec{A}$-dependence of $\{|i\rangle\}_{i=1}^{D_H}$ makes things even more complicated (some further details of this subject are provided in Appendix A, Subsection 11.1).

4.3.4 Complete treatment: method of Lagrange multipliers

The above discussion shows, that in the case of the MF models, MaxEnt principle cannot be applied exactly in a way it is applied in the standard statistical mechanics. We have to search for the probability distribution that maximizes (4.10), and which additionally fulfills the self-consistency requirements expressed by Eqs. (4.8). However, the implicit dependence of the mean-field parameters on probabilities makes the problem nontrivial. In consequence, here we take an alternative route. We postulate that (4.9) should be maximized also with respect to mean-fields $\vec{A}$. Such an approach is motivated by both the Hartree-Fock MF approach, as well as by the Landau theory of phase transitions.

In order to preserve self-consistency of the approach, we make use of the method of Lagrange multipliers. Namely, we supplement (4.9) with the constraint term of the form

$$-\beta \mathrm{Tr}[\hat{Q}_\lambda] = \sum_{s=1}^{M} \beta\lambda_s (\mathrm{Tr}[\hat{\rho}_\lambda \hat{A}_s] - A_s). \tag{4.13}$$

This step indeed ensures that once the constraints are fulfilled, the components of $\vec{A}$, i.e., $A_1, \ldots, A_M$, are the average values of the corresponding operators $\hat{A}_1, \ldots, \hat{A}_M$. At the same time, it allows us to treat D_H diagonal matrix elements of $\hat{\rho}_\lambda$, M mean-fields A_s, and M just introduced Lagrange multipliers λ_s as *independent* variables. In (4.13), $\hat{\rho}_\lambda$ is a new MF density

operator, its form will be subsequently determined with the help the following functional

$$\mathcal{S}_\lambda = \mathrm{Tr}\Big(- \hat{\rho}_\lambda \ln \hat{\rho}_\lambda - \beta \hat{\rho}_\lambda \hat{H} + \sum_{s=1}^{M} \beta \lambda_s \hat{\rho}_\lambda (\hat{A}_s - A_s) + \beta \mu \hat{\rho}_\lambda \hat{N} - \omega\big(\hat{\rho}_\lambda - \frac{1}{D_H}\big)\Big). \tag{4.14}$$

Note, that (4.14) has a canonical form, i.e., all the constraints are linear in $\hat{\rho}_\lambda$. Also, introduction of additional constraints (4.13) may be regarded as a redefinition of the original MF Hamiltonian $\hat{H}$ according to the prescription

$$\hat{H} \;\rightarrow\; \hat{H}_\lambda = \hat{H} + \hat{Q}_\lambda = \hat{H} - \sum_{s=1}^{M} \lambda_s (\hat{A}_s - A_s). \tag{4.15}$$

For $\hat{H}_\lambda$ (4.15) we define also the corresponding grand Hamiltonian $\hat{K}_\lambda$,

$$\hat{K}_\lambda \equiv \hat{H}_\lambda - \mu \hat{N}. \tag{4.16}$$

From (4.14) and (4.15) we infer that $\hat{K}_\lambda$ (4.16), and not $\hat{K} \equiv \hat{H} - \mu \hat{N}$ should be regarded as a true MF grand Hamiltonian of the system.

4.3.5 Trivial time dependence of equilibrium mean-field Hamiltonian and density operator

MF density operator $\hat{\rho}_\lambda$ appearing in (4.14), as well as mean-fields and Lagrange multipliers should be constants of motion for the equilibrium situation, i.e., for the optimal values of $\vec{A}$ and $\vec{\lambda}$ (denoted from now on as $\vec{A} = \vec{A}_0$ and $\vec{\lambda} = \vec{\lambda}_0$). Note, that at the very beginning we have assumed that neither $\hat{A}_s$ operators, nor the MF Hamiltonian depend explicitly on time. Additionally, in equilibrium we must have $\partial \vec{A}_0 / \partial t = 0$, $\partial \vec{\lambda}_0 / \partial t = 0$. In such case, from Eq. (4.5), with $\hat{\rho}_e \rightarrow \hat{\rho}_\lambda(\vec{A}_0, \vec{\lambda}_0; t)$ and $\hat{H}_e \rightarrow \hat{K}_\lambda(\vec{A}_0, \vec{\lambda}_0; t)$, it follows that

$$i\hbar \frac{\partial}{\partial t} \hat{\rho}_\lambda(\vec{A}_0, \vec{\lambda}_0; t) = [\hat{K}_\lambda(\vec{A}_0, \vec{\lambda}_0; t), \hat{\rho}_\lambda(\vec{A}_0, \vec{\lambda}_0; t)] = 0. \tag{4.17}$$

Therefore, the equilibrium (i.e., true grand-canonical) density operator $\hat{\rho}_\lambda(\vec{A}_0, \vec{\lambda}_0; t)$ commutes with the corresponding equilibrium MF grand Hamiltonian $\hat{K}_\lambda(\vec{A}_0, \vec{\lambda}_0; t)$.

However, here we make much stronger assumption, which is crucial for the subsequent construction of the present formalism. Namely, we assume that for *any* initial values of $\vec{A}$ and $\vec{\lambda}$ and for any $s = 1, \ldots, M$ we have

$$\frac{\partial \langle \hat{A}_s \rangle_t}{\partial t} = \frac{\partial A_s(t)}{\partial t} = 0, \qquad \frac{\partial \lambda_s(t)}{\partial t} = 0; \qquad [\hat{K}_\lambda(\vec{A}, \vec{\lambda}; t), \hat{\rho}_\lambda(\vec{A}, \vec{\lambda}; t)] = 0. \tag{4.18}$$

In other words, classical variables $\vec{A}$ and $\vec{\lambda}$ play role of state labels, and have no dynamics (in Sub-subsection 4.11.1 we argue, that such assumption is indeed internally consistent). As a consequence of (4.18), for each $\vec{A}$, $\vec{\lambda}$ we have

$$i\hbar \frac{\partial}{\partial t} \hat{\rho}_\lambda(\vec{A}, \vec{\lambda}; t) = [\hat{K}_\lambda(\vec{A}, \vec{\lambda}; t), \hat{\rho}_\lambda(\vec{A}, \vec{\lambda}; t)] = 0. \tag{4.19}$$

From (4.18) and (4.19) it follows that for each $\vec{A}$ and $\vec{\lambda}$ the we deal with steady state, i.e., the mean field variables, the MF grand Hamiltonian $\hat{K}_\lambda(\vec{A}, \vec{\lambda})$ and the density operator $\hat{\rho}_\lambda(\vec{A}, \vec{\lambda})$ are time-independent. Therefore, in what follows, (trivial) time dependence those quantities will be ignored.

In particular, the condition (4.19) is satisfied for any $\hat{\rho}_\lambda$ which is an operator-valued function of $\hat{K}_\lambda$. Note, that in general we have $[\hat{H}-\mu\hat{N},\hat{\rho}_\lambda]\neq 0$, as well as $[\hat{H}_\lambda,\hat{N}]\neq 0$. Nonetheless, Eq. (4.19) guarantees, the for any $\vec{A}$ and $\vec{\lambda}$ there exist common eigenbasis of $\hat{K}_\lambda$ and $\hat{\rho}_\lambda$, i.e., $\{|i(\vec{A},\vec{\lambda})\rangle\}_{i=1}^{D}$.

In order to make our attempt to describe the system solely in terms of the MF (grand) Hamiltonian $\hat{K}_\lambda(\vec{A},\vec{\lambda})$ complete, a remark is in place here. Namely, from the assumptions made above, it follows that the evolution of the pure MF quantum states ($T=0$) is given by

$$|\Psi(t)\rangle = e^{-it\hat{K}_\lambda(\vec{A},\vec{\lambda})/\hbar}|\Psi(0)\rangle. \tag{4.20}$$

This means, that the time-dependent Schrödinger equation should be always separable into time-independent equation, and that the time dependence of the eigenvectors $|i(\vec{A},\vec{\lambda})\rangle$ of $\hat{K}_\lambda(\vec{A},\vec{\lambda})$ is trivial. Within the present approach, it is not legitimate to analyze non-trivially time-dependent situation, i.e., time evolution of some general state $|\Psi(t)\rangle$, not being an eigenvector of $\hat{K}_\lambda(\vec{A},\vec{\lambda})$ as this would contradict conditions (4.18). This is also discussed in Subsection 4.11.

4.4 Explicit form of mean-field density operator and the optimal (equilibrium) values of mean fields

In order to obtain an explicit form of the grand-canonical MF density operator, we have to find a maximum of $\mathcal{S}_\lambda$ (4.14) subject to the constraints. First, we rewrite (4.14) using the eigenbasis of $\hat{K}_\lambda$, $\{|i(\vec{A},\vec{\lambda})\rangle\}_{i=1}^{D_H}$ (from now on for simplicity denoted $\{|i\rangle\}_{i=1}^{D_H}$), i.e.,

$$\mathcal{S}_\lambda = \sum_{i=1}^{D_H}\Big\{ -q_i\ln q_i - \beta q_i\Big([\hat{H}]_{ii}-\mu[\hat{N}]_{ii}-\sum_{s=1}^{M}\lambda_s\big([\hat{A}_s]_{ii}-A_s\big)\Big) - \omega\Big(q_i-\frac{1}{D_H}\Big)\Big\}. \tag{4.21}$$

In the above, we have $[\hat{N}]_{ii}=\langle i|\hat{N}|i\rangle$ etc., and $q_i=[\hat{\rho}_\lambda]_{ii}$. Also, we have

$$[\hat{H}]_{ii}-\mu[\hat{N}]_{ii}-\sum_{s=1}^{M}\lambda_s\big([\hat{A}_s]_{ii}-A_s\big)=K_\lambda(i). \tag{4.22}$$

where $K_\lambda(i)$ is an eigenvalue of $\hat{K}_\lambda$, i.e., $\hat{K}_\lambda|i\rangle=K_\lambda(i)|i\rangle$.

Next, we assume, that at least one maximum of (4.21) subject to the constraints exists, and that it corresponds to a vanishing gradient of $\mathcal{S}_\lambda$. Then, the necessary conditions for such maximum of $\mathcal{S}_\lambda$ are the following. First,

$$\frac{\partial\mathcal{S}_\lambda}{\partial\omega}=1-\sum_{j=1}^{D_H}q_j=0, \tag{4.23}$$

is the normalization condition. Next, for each $j=1,2,\ldots,D_H$ we have

$$\frac{\partial\mathcal{S}_\lambda}{\partial q_j}=-(1+\omega)-\ln q_j-\beta[\hat{H}]_{jj}+\beta\mu[\hat{N}]_{jj}+\sum_{s=1}^{M}\beta\lambda_s\big([\hat{A}_s]_{jj}-A_s\big)=0. \tag{4.24}$$

Also, for each $w=1,2,\ldots,M$ we have

$$\frac{\partial\mathcal{S}_\lambda}{\partial A_w}=-\beta\sum_{i=1}^{D_H}q_i\left(\frac{\partial[\hat{H}]_{ii}}{\partial A_w}+\lambda_w\right)=0, \tag{4.25}$$

and

$$\frac{\partial \mathcal{S}_\lambda}{\partial \lambda_w} = \beta \sum_{i=1}^{D_H} q_i \big([\hat{A}_w]_{ii} - A_w \big) = 0. \tag{4.26}$$

In (4.25) and (4.26), in the terms $[\hat{K}_\lambda]_{ii} = \langle i|\hat{K}_\lambda|i\rangle = K_\lambda(i)$ we have ignored the possible explicit $\vec{A}$- and $\vec{\lambda}$- dependence of the (normalized) eigenvectors $|i\rangle$. This is justified due to the Hellmann-Feynman theorem [180]. Also, let us point out once more that due to the presence of Lagrange multipliers λ_s, the averages A_s (mean-fields), and probabilities $q_i = \langle i|\hat{\rho}_\lambda|i\rangle$ are treated here as independent variables.

4.4.1 Variational parameters of a non- mean-field character

Apart from the mean-fields $\vec{A}$, mean-field Hamiltonian $\hat{H}(\vec{A})$ may depend also on some extra variables $(b_1, b_2, \ldots, b_P) \equiv \vec{b}$, which are not expectation values of operators of the same type as those used to construct the MF Hamiltonian. For example, for the lattice fermion systems, this means that b_l are not averages of expressions bilinear in fermion creation or annihilation operators. A lattice (bond) lengths or variational parameters originating from Gutzwiller [20] or Jastrow [33, 181] correlators present in trial wave functions may serve as examples. Certain parameters of the MF Hamiltonian, e.g. hopping integrals between more distant neighbors may be also treated as variational parameters of the same type. If values of $b_1, b_2, \ldots, b_P$ are not *a priori* known, and are to be predicted within our model, we may treat them as additional variational parameters, with respect to which $\mathcal{S}_\lambda$ is maximized. Clearly, no corresponding Lagrange multipliers are required in such case. Explicitly, if the maximum of $\mathcal{S}_\lambda$ with respect to each b_l corresponds to a stationary point[21] (i.e., point of vanishing gradient of $\mathcal{S}_\lambda$), using Eq. (4.21) we obtain, apart from Eqs. (4.23)-(4.26), P additional necessary conditions of the form

$$\frac{1}{\beta}\frac{\partial \mathcal{S}_\lambda}{\partial b_l} = -\sum_{i=1}^{D_H} q_i \frac{\partial [\hat{H}]_{ii}}{\partial b_l} = 0. \tag{4.27}$$

The last equality holds as long as the constraint part $\hat{Q}_\lambda$ of $\hat{H}_\lambda$ does not depend on b_l, which is usually the case.

Note, that maximization of (4.14) with respect to $\vec{b}$ variables leads us beyond the standard formulation of the MaxEnt principle, where only the probabilities play role of variational parameters. Nonetheless, such an extension seems to be legitimate. According to the MaxEnt philosophy, we look for the maximum of S_{vN} (4.4) prior to constraints of given $E = \langle \hat{H} \rangle$ and $N = \langle \hat{N} \rangle$. This procedure leads to the one-to-one correspondence between E and β. Alternatively, we work with fixed β and variable E. In any case, the same necessary conditions (4.27) for the extremum follow.

Also, please note, that maximization of $\mathcal{S}_\lambda$ with respect to $\vec{b}$ is equivalent to the more familiar procedure, namely to minimization of the corresponding generalized grand potential. It is well-known, that for given T, V, μ and other thermodynamic variables, e.g. external electric or magnetic field, in the equilibrium situation the thermodynamic grand potential Ω (4.47) reaches its minimal value. In the $\beta \to \infty$ limit, the entropic part of $\mathcal{S}_\lambda$ becomes of no importance and we are left with minimization of the ground-state energy, i.e., the variational principle of quantum mechanics for the MF (grand) Hamiltonian.

[21] Obviously, this may be not the case. As an example, we may invoke MF models describing Fulde-Ferrel-Larkin- Ovchinnikov (FFLO) type of superconducting state (cf. e.g. Refs. [182, 183] and References therein). The corresponding MF Hamiltonians depend on the center of mass wave vector $\vec{q}$ of the Cooper pairs, which appear as a quantum number labeling creation and annihilation operators. In such case it would be very unnatural to employ the conditions (4.27). Rather, we have to examine each value of $\vec{q}$ separately and select the one corresponding to maximum of $\mathcal{S}_\lambda$.

$\mathcal{S}_\lambda$ (4.14) depends on parameters $\vec{b}$ only through the MF Hamiltonian, similarly to the dependencies of $\mathcal{S}_\lambda$ on volume V or an external (electric or magnetic) field $\vec{h}$. However, in contrast to V or $\vec{h}$, values of $b_1, b_2, \ldots$ are not specified until the model is solved (one can think of a family of MF Hamiltonians labeled by $\vec{b}$). Note, that in general, $[\hat{H}(\vec{b}_1), \hat{H}(\vec{b}_2)] \neq 0$ for $\vec{b}_1 \neq \vec{b}_2$.

On the other hand, if the value of some of $b_1, \ldots, b_P$ parameters (say, of b_l) becomes known (which is the case for the standard parameters of the Hamiltonian, like particle mass, Coulomb or hopping integrals, external fields etc.), we obviously no longer treat b_l as a variational parameter. Then Eq. (4.27) is not obeyed, and we obtained lower value of $\mathcal{S}_\lambda$, i.e., more informative probability distribution. In such case we change the notation, i.e., $b_l \to \alpha_l$ as in Ref. [109]. The following quantity

$$\phi_{l0} = \left\langle \frac{\partial \hat{H}}{\partial \alpha_l} \right\rangle_0 = -\frac{1}{\beta} \left(\frac{\partial \mathcal{S}}{\partial \alpha_l} \right)_0 \tag{4.28}$$

corresponds to the optimal MaxEnt estimate of $\phi_l \equiv \langle \partial \hat{H}/\partial \alpha_l \rangle$. In the above, subscript '0' indicates, that we use values of p_i, A_s and λ_s which are the solution of Eqs.(4.23)-(4.26). In that way we may determine the value of pressure (then $\alpha_l = V$, the volume of a specimen) or the magnetic moment (magnetization) with $\alpha_l = \vec{h}\hat{z}$, the z-the component of the external magnetic field. If parameters $\vec{b}$ are present in a problem, formulas (4.1), (4.2) and (4.8)-(4.26) may be easily generalized.

4.4.2 Observables with *a priori* known expectation values

Apart from $N = \langle \hat{N} \rangle \equiv \langle \hat{A}_1 \rangle$, expectation values of some other $\hat{A}_s$ operators may be *a priori* known (for simplicity, assume that we have one such operator, say $\hat{A}_2$). In that situation, we do not insert the numerical value of $\langle \hat{A}_2 \rangle$ into $\hat{H}(\vec{A})$. Instead, in order to ensure that the value of $\langle \hat{A}_2 \rangle$ is consistent with the prior information we have, we add to $\hat{K}_\lambda$ the term $(-\xi_2 \hat{A}_2)$ (analogous to the $(-\mu \hat{N})$ term). If we ignore some available information, we obtain less informative probability distribution. Note, that even if its value becomes known, $A_2 = \langle \hat{A}_2 \rangle$ is still treated as a variational parameter (see discussion in Subsection 4.7) and $\mathcal{S}_\lambda$ has to be maximized with respect to all variables A_s. Otherwise, the standard thermodynamic relations (cf. Section 4.6) would not be valid. This is in contrast to the case of b_l parameters, discussed in the previous Sub-subsection. The asymmetric treatment of those two types of variables is simply caused by the fact, that for the b_l variables there are no corresponding operators.

4.4.3 Explicit form of functional dependence of mean-field density operator on $\vec{A}, \vec{\lambda}$ and $\vec{b}$ variables

Eqs. (4.23) and (4.24) can be easily solved, and q_i and ω may be expressed in terms of $\vec{A}$, $\vec{\lambda}$ and $\vec{b}$.[22] Together with the condition (4.19), this yields a basis-independent form of the MF density operator $\hat{\rho}_\lambda(\vec{A}, \vec{\lambda}, \vec{b})$ and the corresponding partition function $\mathcal{Z}_\lambda(\vec{A}, \vec{\lambda}, \vec{b})$, namely

$$\begin{aligned} \hat{\rho}_\lambda(\vec{A}, \vec{\lambda}, \vec{b}) &= \mathcal{Z}_\lambda^{-1} \exp\big(-\beta(\hat{H}_\lambda(\vec{A}, \vec{\lambda}, \vec{b}) - \mu \hat{N})\big), \\ \mathcal{Z}_\lambda(\vec{A}, \vec{\lambda}, \vec{b}) &= \mathrm{Tr}[\exp\big(-\beta(\hat{H}_\lambda(\vec{A}, \vec{\lambda}, \vec{b}) - \mu \hat{N})\big)]. \end{aligned} \tag{4.29}$$

Note, that without the assumption (4.19), which allows to match the algebraic structure of the Hilbert space with Eqs. (4.23) and (4.24), the above form of $\hat{\rho}_\lambda$ could not be obtained. Next,

[22] The latter variables are explicitly present only in some of the following formulas.

we rewrite Eqs. (4.25) and (4.26) as

$$-\frac{1}{\beta}\frac{\partial \mathcal{S}_\lambda}{\partial A_w} = \mathrm{Tr}\left[\hat{\rho}_\lambda\Big(\frac{\partial \hat{H}}{\partial A_w} + \lambda_w\Big)\right] \equiv \left\langle \frac{\partial \hat{H}}{\partial A_w} \right\rangle + \lambda_w = \left\langle \frac{\partial \hat{H}_\lambda}{\partial A_w} \right\rangle = 0, \tag{4.30}$$

$$\frac{1}{\beta}\frac{\partial \mathcal{S}_\lambda}{\partial \lambda_w} = \mathrm{Tr}[\hat{\rho}_\lambda(\hat{A}_w - A_w)] \equiv \left\langle \hat{A}_w \right\rangle - A_w = \left\langle \frac{\partial \hat{H}_\lambda}{\partial \lambda_w} \right\rangle = 0. \tag{4.31}$$

Similarly, Eqs. (4.27) can be given the form

$$-\frac{1}{\beta}\frac{\partial \mathcal{S}_\lambda}{\partial b_l} = \mathrm{Tr}\left[\hat{\rho}_\lambda \frac{\partial \hat{H}}{\partial b_l}\right] \equiv \left\langle \frac{\partial \hat{H}}{\partial b_l} \right\rangle = \left\langle \frac{\partial \hat{H}_\lambda}{\partial b_l} \right\rangle = 0. \tag{4.32}$$

In the above, and in what follows by $\langle \ldots \rangle$ we understand[23] $\mathrm{Tr}[\hat{\rho}_\lambda(\ldots)]$.

4.4.4 Generalized grand potential

We define the *generalized grand potential* as

$$\mathcal{F}(\vec{A}, \vec{\lambda}, \vec{b}) \equiv -\beta^{-1} \ln \mathcal{Z}_\lambda(\vec{A}, \vec{\lambda}, \vec{b}) \equiv -\beta^{-1}\mathcal{S}_\lambda(\vec{q}(\vec{A}, \vec{\lambda}, \vec{b}), \vec{A}, \vec{\lambda}, \vec{b}), \tag{4.33}$$

with $\mathcal{Z}_\lambda$ given by (4.29). Note, that $\mathcal{F}(\vec{A}, \vec{\lambda}, \vec{b})$ is a function of the standard thermodynamic variables $T, \mu, V, \ldots$, the mean-field variables $\vec{A}$, $\vec{\lambda}$ and other variables of non-MF character,[24] i.e., $\vec{b}$. From now on we will work with $\mathcal{F}$ rather then with $\mathcal{S}_\lambda$. Equations (4.30)-(4.32) may be thus given the form

$$\nabla_A \mathcal{F} = \vec{0}_M, \qquad \nabla_\lambda \mathcal{F} = \vec{0}_M, \quad \text{and} \quad \nabla_b \mathcal{F} = \vec{0}_P. \tag{4.34}$$

In the above, we have $\nabla_A \equiv \big(\frac{\partial}{\partial A_1}, \frac{\partial}{\partial A_2}, \ldots, \frac{\partial}{\partial A_M}\big)$, and analogously for ∇_λ and ∇_b. $\vec{0}_M$ ($\vec{0}_P$) denotes the M- (P-) dimensional zero vector, i.e., $\vec{0} = (0, 0, \ldots, 0)$. Despite the fact, that derivatives of $\hat{K}_\lambda$ with respect to A_s, λ_s, and b_l does not have to commute with $\hat{K}_\lambda$, Eqs. (4.34) and (4.30)-(4.32) are equivalent. This is a consequence of the following identity (cf. [178])

$$\frac{\partial}{\partial x} e^{\hat{C}(x)} = \int_0^1 d\tau e^{\tau \hat{C}(x)} \Big(\frac{\partial}{\partial x}\hat{C}(x)\Big) e^{(1-\tau)\hat{C}(x)}, \tag{4.35}$$

where $\hat{C}(x)$ is an arbitrary operator, as well as from invariance of trace with respect to cyclic permutation of the operators.[25]

The form of Eqs. (4.30)-(4.32) and (4.34) guarantees that they are valid regardless the particular choice of the basis states. It is usually most convenient to work in the eigenbasis of $\hat{K}_\lambda$, but sometimes an $\vec{A}$, $\vec{\lambda}$-independent basis may be more useful.

In contrast to the non-MF case, where the MaxEnt principle yields a unique solution, by solving (4.34), for given $\beta, \mu, V, \ldots$ (or given the number of lattice sites Λ instead of the volume V) and parameters of the Hamiltonian, we usually obtain a finite number[26] of stationary points

[23] Note that this corresponds to $\langle \ldots \rangle_\lambda$ in the notation of Ref. [113].

[24] within the notation for $\mathcal{F}$ we use here, both the thermodynamic parameters, e.g. temperature T or volume V, as well as the b_l variables of Sub-subsection 4.4.1 are usually omitted, i.e., we write $\mathcal{F}(\vec{A}, \vec{\lambda})$.

[25] Instead of using Eq. (4.35), one may also expand exponents of (4.29) in the Taylor series, compute derivative, and then make use of the properties of trace operation with respect to each such obtained term.

[26] This may be not the case e.g. for the superconducting order parameter, represented by a complex number $A_s + iA_{s+1} = \Delta = |\Delta|e^{i\varphi}$. Then, value of $\mathcal{F}$ does not depend on φ, which expresses unbroken $U(1)$ symmetry. Therefore, there exist infinitely many solutions parameterized by φ ('flat direction'). Nonetheless, we may restrict ourselves to real Δ from the beginning, and then only a simple degeneracy with respect to sign of an order parameter is present.

of $\mathcal{F}$. Nonetheless, some of them may not correspond to the physically relevant situations, as will be discussed later.

Next, we select the solution of (4.34), which is characterized by the lowest value of $\mathcal{F}$. However, in the absence of external, symmetry-breaking fields, such solution is usually not unique. In that situation we explicitly break the existing symmetry 'by hand', i.e., among the equivalent minima we select the one characterized by well-defined values of the mean-fields, $\vec{A} \equiv \vec{A}_0$, Lagrange multipliers $\vec{\lambda} \equiv \vec{\lambda}_0$, and other variational parameters, $\vec{b} \equiv \vec{b}_0$, which corresponds to thermodynamic equilibrium. For the most cases of interest, Eqs. (4.34) are not analytically soluble and we must rely on numerical analysis (see Part III for concrete applications).

Moreover, Eqs. (4.34) provide only the necessary, not the sufficient conditions for the existence of the minimum of $\mathcal{F}$ (maximum of $\mathcal{S}_\lambda$ (4.14)). Note, that in contrast to the non-MF case, here the stationary points of $\mathcal{F}$ do not have to correspond to the minimum. The explicit analytical conditions for the existence of maximum of $\mathcal{S}_\lambda$ prior to constraints can be provided, but in practice it is more convenient to analyze the character of each stationary point separately.

4.4.5 Grand-canonical (equilibrium) mean-field Hamiltonian and density operator

True grand-canonical (equilibrium) mean-field grand Hamiltonian, invoked in Sub-subsection 4.3.5, may be now explicitly defined as

$$\hat{K}_{\lambda 0} = \hat{K}_\lambda(\vec{A}_0, \vec{\lambda}_0, \vec{b}_0) = \hat{H}_\lambda(\vec{A}_0, \vec{\lambda}_0, \vec{b}_0) - \mu\hat{N}. \tag{4.36}$$

Consequently, grand-canonical MF density operator, denoted $\hat{\rho}_{\lambda 0}$, is now given by Eq. (4.29) with $\vec{A} = \vec{A}_0, \vec{\lambda} = \vec{\lambda}_0$, and $\vec{b} = \vec{b}_0$, i.e.,

$$\hat{\rho}_{\lambda 0} = \mathcal{Z}_{\lambda 0}^{-1} \exp(-\beta \hat{K}_{\lambda 0}), \qquad \mathcal{Z}_{\lambda 0} = \mathrm{Tr}[\exp(-\beta \hat{K}_{\lambda 0})]. \tag{4.37}$$

Due to the particular choice of the Hamiltonian parameters, $\hat{H}(\vec{A})$ may cease to depend on some mean field, say $A_s \in \{A_1, \ldots, A_M\}$. Than, from (4.30) it follows that the corresponding Lagrange multiplier vanishes in the equilibrium situation, $\lambda_{s0} = 0$. On the other hand, it is possible that $\mathrm{Tr}[\hat{A}_s \hat{\rho}_{\lambda 0}] \neq 0$, even that $\hat{A}_s$ does not appear in $\hat{H}(\vec{A})$.

4.4.6 Approach based solely on Bogoliubov-de Gennes self-consistent equations

The results of a non-variational self-consistent approach based on BdG equations (4.8) may be obtained easily within the present formalism. Namely, by equating the derivatives of $\mathcal{F}$ with respect to all λ_s to zero, i.e., $\nabla_\lambda \mathcal{F} = 0$, and then subsequently putting $\vec{\lambda} = \vec{0}$, we recover Eqs. (4.8) with MF density operator $\hat{\rho}(\vec{A}) \equiv \hat{\rho}_\lambda(\vec{A}, \vec{\lambda} = \vec{0})$. If additional variational parameters of non-mean field character, $b_1, b_2, \ldots, b_P$ are present, we minimize $\mathcal{F}(\vec{A}, \vec{0}, \vec{b})$ with respect to the latter variables, in order to find their optimal values (cf. the corresponding Eqs. (4.27) for the MaxEnt-based fully variational approach). In such a situation, we have to solve the following set of equations

$$\nabla_\lambda \mathcal{F}(\vec{A}, \vec{0}, \vec{b}) = \vec{0}_M, \qquad \nabla_b \mathcal{F}(\vec{A}, \vec{0}, \vec{b}) = \vec{0}_P. \tag{4.38}$$

The optimal solution of Eqs. (4.38), i.e., $\vec{A}$ and $\vec{b}$, for which $\mathcal{F}(\vec{A}, \vec{0}, \vec{b})$ has a minimal value, will be denoted $\vec{A}_{sc}^{(0)}$ and $\vec{b}_{sc}^{(0)}$, respectively. Therefore, the optimal MF (grand) Hamiltonian, corresponding to $\hat{K}_{\lambda 0}$ (4.36) of the variational approach reads

$$\hat{K}_{sc}^{(0)} = \hat{K}_\lambda(\vec{A}_{sc}^{(0)}, \vec{0}, \vec{b}_{sc}^{(0)}) = \hat{H}(\vec{A}_{sc}^{(0)}, \vec{b}_{sc}^{(0)}) - \mu\hat{N}. \tag{4.39}$$

The optimal MF density operator, corresponding to $\hat{\rho}_{\lambda 0}$ (4.37), has now the following form

$$\hat{\rho}_{sc}^{(0)} = \left(\mathcal{Z}_{sc}^{(0)}\right)^{-1} \exp(-\beta \hat{K}_{sc}^{(0)}), \qquad \mathcal{Z}_{sc}^{(0)} = \mathrm{Tr}[\exp(-\beta \hat{K}_{sc}^{(0)})]. \tag{4.40}$$

Note, that $\hat{\rho}_{sc}^{(0)} = \hat{\rho}(\vec{A}_{sc}^{(0)}, \vec{b}_{sc}^{(0)})$ for $\hat{\rho}(\vec{A}, \vec{b})$ given by Eq. (4.11).[27]

4.5 Non-equilibrium situation and relation of the present approach to Landau theory of phase transitions

4.5.1 Self-consistency conditions for arbitrary values of mean-fields

So far, our aim has been to construct the MF density operator (4.29), and, by solving Eqs. (4.34), to obtain the optimal values of mean-fields ($\vec{A} = \vec{A}_0$), the Lagrange multipliers ($\vec{\lambda} = \vec{\lambda}_0$), and additional variational parameters ($\vec{b} = \vec{b}_0$), which correspond to the equilibrium situation. Here, we will consider a non-equilibrium situation, i.e., $\mathcal{D}_{\mathcal{A}} \ni \vec{A} \neq \vec{A}_0$.[28]

Within any MF approach without the constraint terms (4.13) present in the MF Hamiltonian, a set of permitted values of mean-fields $\mathcal{D}_{\mathcal{A}}$ reduces usually to a small number of isolated points, $\vec{A} \in \{\vec{A}_{sc}^{(1)}, \vec{A}_{sc}^{(2)}, \ldots, \vec{A}_{sc}^{(R)}\}$, determined by solutions of the self-consistency equations (4.8). In contrast, within the present approach it is fully legitimate to consider any $\vec{A} \in \mathcal{D}_{\mathcal{A}}$. In order to do that, for each $\vec{A}$, values of $\vec{\lambda}$ have to be chosen in a way that guarantees, that the self-consistency conditions (4.8) are fulfilled for each point $\vec{A}$. Explicitly, we have to solve Eqs. (4.8) or, equivalently, M of equations (4.34), i.e.,

$$\frac{\partial \mathcal{F}}{\partial \lambda_1} = 0, \qquad \frac{\partial \mathcal{F}}{\partial \lambda_2} = 0, \ldots, \qquad \frac{\partial \mathcal{F}}{\partial \lambda_M} = 0. \tag{4.41}$$

Note, that Eqs. (4.41) may have no solution. This implies that the particular value of $\vec{A}$ cannot be obtained consistently within a model under consideration, and consequently, $\vec{A} \notin \mathcal{D}_{\mathcal{A}}$.[29] Next, Eqs. (4.41) may have an unique solution. Finally, there may exist p different solutions, $\vec{\lambda}_1(\vec{A}), \ldots, \vec{\lambda}_p(\vec{A})$. In such case, we select this solution (denoted by $\vec{\lambda}(\vec{A})$), for which $\mathcal{F}$ has the lowest value. In the two latter cases, for each $\vec{A} \in \mathcal{D}_{\mathcal{A}}$ we can define a *self-consistent* MF Hamiltonian

$$\hat{H}_z(\vec{A}) \equiv \hat{H}_\lambda(\vec{A}, \vec{\lambda}(\vec{A})), \tag{4.42}$$

the corresponding density operator,

$$\hat{\rho}_z(\vec{A}) = \hat{\rho}_\lambda(\vec{A}, \vec{\lambda}(\vec{A})) = \mathcal{Z}_z^{-1} e^{-\beta(\hat{H}_z - \mu\hat{N})}, \qquad \mathcal{Z}_z = \mathrm{Tr}[e^{-\beta(\hat{H}_z - \mu\hat{N})}], \tag{4.43}$$

and the self-consistent grand potential

$$\mathcal{F}_z(\vec{A}) \equiv \mathcal{F}(\vec{A}, \vec{\lambda}(\vec{A})) = -\beta^{-1} \ln \mathcal{Z}_z(\vec{A}). \tag{4.44}$$

Note, that in order to obtain $\mathcal{F}_z(\vec{A})$ and $\hat{\rho}_z(\vec{A})$, we have to maximize $\mathcal{S}_\lambda$ with respect to the probabilities only, with the values of all mean-fields being fixed. Therefore, for a given $\hat{H}(\vec{A})$, $\hat{\rho}_z(\vec{A})$ represents an optimal probability distribution prior to M constraints $\langle \hat{A}_s \rangle = A_s$ (which expressed in terms of the generalized grand potential $\mathcal{F}(\vec{A}, \vec{\lambda})$ are given by Eqs. (4.41). The latter conditions determine values of the Lagrange multipliers.

$\mathcal{F}_z(\vec{A})$ depends both on the standard thermodynamic variables $T, V, \mu, \vec{h}, \ldots$ (by $\vec{h}$ we denote the external either magnetic or electric field), as well as on the mean-fields $\vec{A}$.[30] For

[27] $\vec{b}$ variables are not explicitly present in (4.11).

[28] For simplicity, in what follows, the variational parameters $\vec{b}$ will be omitted in most formulas. They can be easily reintroduced, if necessary.

[29] Apart from $\mathcal{D}_{\mathcal{A}}$, i.e., set of those $\vec{A}$, for which $\hat{H}(\vec{A})$ is well-defined, one may define also another set, $\mathcal{D}_{\mathcal{A}}^*$, by the condition that if $\vec{A} \in \mathcal{D}_{\mathcal{A}}^*$, then at least one solution of Eqs. (4.41) exists. Obviously, $\mathcal{D}_{\mathcal{A}}^* \subseteq \mathcal{D}_{\mathcal{A}}$ and usually $\mathcal{D}_{\mathcal{A}}^* \neq \mathcal{D}_{\mathcal{A}}$, but we will not distinguish between those two sets.

[30] For a given N, chemical potential μ is determined by Eqs. (4.34), cf. Sec. 4.7.

simplicity, the former variables, as well as e.g. the number of the lattice sites Λ are omitted, in analogy with the notation for $\mathcal{F}(\vec{A},\vec{\lambda})$ in the previous Section. More precise notation, namely $\mathcal{F}_z^{(\Lambda)}(T,V,\mu,\vec{h};\vec{A})$ or $\mathcal{F}_z^{(\Lambda)}(T,V,\mu,\vec{h};\vec{A},\vec{b})$ will be used only when the distinction between those two kinds of variables is crucial, e.g. in Sec. 4.6.

4.5.2 Interpretation of $\mathcal{F}_z(\vec{A})$ as Landau potential

Some of the mean-fields have usually a natural interpretation of order parameters (e.g. superconducting gap magnitude, magnetization, etc.). Therefore it is tempting to identify $\mathcal{F}_z(\vec{A})$ (4.44) with the Landau (grand) potential [148, 149] related to a given microscopic MF Hamiltonian $\hat{H}(\vec{A})$. We argue, that such identification is justified and hence we may try to incorporate the thermodynamic fluctuation into a generalized Landau scheme (details are provided in Supplement A, Subsection 12.1). First, note that we have

$$\frac{\partial \mathcal{F}_z(\vec{A})}{\partial A_s} = \frac{\partial \mathcal{F}(\vec{A},\vec{\lambda}(\vec{A}))}{\partial A_s} + \sum_{t=1}^{M} \frac{\partial \mathcal{F}(\vec{A},\vec{\lambda}(\vec{A}))}{\partial \lambda_t} \cdot \frac{\partial \lambda_t(\vec{A})}{\partial A_s}. \tag{4.45}$$

Therefore, if $\partial \mathcal{F}_z(\vec{A})/\partial A_s = 0$, then from Eqs. (4.41) and (4.45) it follows that we also have $\partial \mathcal{F}_\lambda(\vec{A},\vec{\lambda}(\vec{A}))/\partial A_s = 0$, i.e., that Eqs. (4.34) are fulfilled for $\vec{A}$ and $\vec{\lambda} = \vec{\lambda}(\vec{A})$. Conversely, if $\vec{A}$ and $\vec{\lambda} = \vec{\lambda}(\vec{A})$ form a solution of (4.34), we have $\partial \mathcal{F}_z(\vec{A})/\partial A_s = 0$. Then, each stationary point of $\mathcal{F}(\vec{A},\vec{\lambda})$ as given by (4.34) is in a one-to-one correspondence to a stationary point of $\mathcal{F}_z(\vec{A})$. Consequently, one may repeat part of the discussion of the previous Section using $\mathcal{F}_z(\vec{A})$ instead of $\mathcal{F}(\vec{A},\vec{\lambda})$. Namely, only the minima of $\mathcal{F}_z(\vec{A})$ (both global and local) correspond to acceptable solutions (stable and meta-stable, respectively), whereas the maxima and saddle points should be regarded as unphysical. Exactly as in the original Landau approach, the global minimum of $\mathcal{F}_z(\vec{A})$, (which in an absence of the symmetry-breaking terms must be selected 'by hand' out of many equivalent minima) corresponds to the equilibrium situation.

Strictly speaking, we have obtained a more general situation then in the original Landau theory. First, mean-fields which do not have a natural interpretation of order parameters (e.g. average (total) particle number, or average kinetic (band) energy) are usually present in a problem. Also, the original formulation of Landau would be recovered if we expand $\mathcal{F}_z(\vec{A})$ in powers of the relevant order parameters up to the fourth or the sixth order. If a mean-field represents physical quantity invariant with respect to the point-symmetry transformations (i.e., particle density n_i on the given site or sub-lattice), also the odd powers of such order parameter appear in the expansion of $\mathcal{F}_z(\vec{A})$. Yet, all the above mentioned differences are rather unimportant and we will call $\mathcal{F}_z(\vec{A})$ Landau (grand) potential from now on. Unfortunately, the explicit form of $\vec{\lambda}(\vec{A})$, i.e., the analytical solution of (4.41) is rarely available. Consequently, the explicit form of $\mathcal{F}_z(\vec{A})$ is usually not feasible either, and one must rely on numerical analysis (the notable exception is analyzed in Sec. 12.3.)

4.5.3 Incorrect construction of Landau potential

If the self-consistency preserving constraints 4.13 are not present in the MF Hamiltonian, conditions (4.8) are usually not fulfilled for arbitrary $\vec{A}$. As a consequence, such MF model is well-defined only for a finite number of points, $\vec{A} \in \{\vec{A}_{sc}^{(1)}, \vec{A}_{sc}^{(2)}, \ldots, \vec{A}_{sc}^{(R)}\}$. For other values of $\vec{A}$, the model is not internally consistent. However, this fact is sometimes ignored, and it is assumed that the following function of mean fields,

$$\mathcal{F}_n(\vec{A}) \equiv \mathcal{F}(\vec{A},\vec{0}), \tag{4.46}$$

in which the self-consistency requirements are disregarded, can play a role of the Landau potential. Clearly, $\mathcal{F}_n(\vec{A})$ is distinct from $\mathcal{F}_z(\vec{A})$, even for the MF Hamiltonians of the Hartree-Fock form (cf. Eqs (12.300) and (12.302) in Section 12.3). Needless to say, it is not legitimate to use $\mathcal{F}_n(\vec{A})$ instead of $\mathcal{F}_z(\vec{A})$.

4.5.4 Final remarks

Further analysis of the nonequilibrium situation, (with the emphasis on the thermodynamic fluctuations), its relation with the equilibrium situation, as well as further discussion of the properties of Landau potentials constructed within the present approach, are provided in Appendix C (Subsection 11.3) and Supplement A (Subsection 12.1).

We have decided to move large part of this material to the Appendices, because it neither utilized in Part III, nor directly related to its contents. However, the material presented in the Appendices constitutes indispensable part of the whole formalism, and therefore we have decided to include it in this Thesis for the sake of completeness.

4.6 Equilibrium thermodynamics

Now we come back to the analysis of thermodynamic equilibrium and discuss the mean-field thermodynamics. In Subsection 4.4 we have pointed out, that $\vec{A} = \vec{A}_0$, $\vec{\lambda} = \vec{\lambda}_0$ and $\vec{b}_0 = \vec{b}_0$, given by the solution of Eqs. (4.34) characterized by the lowest value of $\mathcal{F}$ (4.33), correspond to the equilibrium situation. Clearly, $\vec{A}_0$, $\vec{\lambda}_0$, and $\vec{b}_0$ usually depend in a non-trivial way on standard thermodynamic variables like temperature T, volume V, chemical potential μ, and the external field, denoted $\vec{h}$.

In Supplement A (Subsection 12.1) we show, that for the most cases of interest, in particular when the number of mean-fields does not depend on Λ, the equilibrium mean-field description is expected to be valid in the thermodynamic limit, i.e., $\Lambda, N, V \to \infty$, $N/\Lambda = n = \text{const}$.

4.6.1 Grand potential

Making use of the generalized grand potential $\mathcal{F}(T, V, \mu, \vec{h}; \vec{A}, \vec{\lambda}, \vec{b})$ or Landau grand potential $\mathcal{F}_z(T, V, \mu, \vec{h}; \vec{A}, \vec{b})$, we proceed in a direct analogy with Landau theory and define true thermodynamic grand potential $\Omega = \Omega(T, V, \mu, \vec{h})$ as

$$\begin{aligned} \Omega(T, V, \mu, \vec{h}) &= \mathcal{F}(T, V, \mu, \vec{h}; \vec{A}_0(T, V, \mu, \vec{h}), \vec{\lambda}_0(T, V, \mu, \vec{h}), \vec{b}_0(T, V, \mu, \vec{h})) \\ &= \mathcal{F}_z(T, V, \mu, \vec{h}; \vec{A}_0(T, V, \mu, \vec{h}), \vec{b}_0(T, V, \mu, \vec{h})). \end{aligned} \tag{4.47}$$

In what follows we use of $\mathcal{F}(T, V, \mu, \vec{h}; \vec{A}, \vec{\lambda}, \vec{b})$ instead of $\mathcal{F}_z(T, V, \mu, \vec{h}; \vec{A}, \vec{b})$, because usually the explicit analytical form of the latter can rarely be obtained.

4.6.2 First derivatives of grand potential

Let $y \in \{T, V, \mu, \vec{h}\}$.[31] We have[32]

$$\frac{\partial \Omega}{\partial y} = \left(\frac{\partial \mathcal{F}}{\partial y}\right)_0 + \sum_s \left\{ \left(\frac{\partial \mathcal{F}}{\partial A_s}\right)_0 \frac{\partial A_{s0}}{\partial y} + \left(\frac{\partial \mathcal{F}}{\partial \lambda_s}\right)_0 \frac{\partial \lambda_{s0}}{\partial y} \right\}. \tag{4.48}$$

[31] By $\vec{h}$ we understand here either the three-dimensional vector external field, or its relevant component. By $\partial\Omega/\partial\vec{h}$ we understand, as usual, a vector of partial derivatives of Ω with respect to the components of $\vec{h}$.

[32] In this and the following formulas we do not include non-MF variational parameters $b_1, \ldots, b_P$ explicitly. As previously, the results provided below can be easily generalized if such variables are present.

In the above, subscript $'0'$ indicates that we insert the equilibrium values and Lagrange multipliers of mean-fields *after* computing the respective derivatives.Terms in curly brackets in Eq. (4.48) vanish due to (4.34), therefore the implicit dependence of Ω on y via $\vec{A}_0(y)$ and $\vec{\lambda}_0(y)$ does not contribute to (4.48), and we recover the same situation as in a standard, non-MF statistical mechanics. This is true, as long as $\hat{H}_\lambda$ does not depend explicitly on T or μ. As a consequence of Eq. (4.48), proper statistical-mechanical expressions for the entropy, pressure, average particle number, magnetization and other quantities given by the first derivatives of Ω are recovered. Note again, that the same situation is encountered in Landau theory. It may be instructive also to show this directly, i.e., by using the corresponding statistical-mechanical definitions. For concreteness, consider $y = T$. We have

$$\begin{aligned}\left(\frac{\partial\Omega}{\partial T}\right)_{V,\mu} &= \beta^2\partial_\beta\big(\beta^{-1}\ln\mathcal{Z}_{\lambda 0}\big) = \beta\mathcal{Z}_{\lambda 0}^{-1}\partial_\beta\mathcal{Z}_{\lambda 0} - \ln\mathcal{Z}_{\lambda 0}\\ &= -\beta\mathcal{Z}_{\lambda 0}^{-1}\mathrm{Tr}[(\beta\partial_\beta\hat{K}_{\lambda 0} + \hat{K}_{\lambda 0})e^{-\beta\hat{K}_{\lambda 0}}] - \ln\mathcal{Z}_{\lambda 0}\\ &= -\beta^2\langle\partial_\beta\hat{K}_{\lambda 0}\rangle - \beta\langle\hat{K}_{\lambda 0}\rangle - \ln\mathcal{Z}_\lambda\\ &= -\beta^2\langle\partial_\beta\hat{K}_{\lambda 0}\rangle + \langle\ln\hat{\rho}_{\lambda 0}\rangle = -\beta^2\langle\partial_\beta\hat{K}_{\lambda 0}\rangle - S. \end{aligned} \tag{4.49}$$

In the above, $\hat{K}_{\lambda 0}$, $\hat{\rho}_{\lambda 0}$, and $\mathcal{Z}_{\lambda 0}$ are given by (4.36) and (4.37). Also, we have $\ln\mathcal{Z}_{\lambda 0} = -\beta\Omega$, and $\langle\hat{K}_{\lambda 0}\rangle = U - \mu N$. Note, that we have implicitly made use of the identity (4.35), as well as linearity of trace and its invariance with respect to cyclic permutation of operators. In effect, we obtain a desired result, provided that the term $\langle\partial_\beta\hat{K}_{\lambda 0}\rangle$ vanishes. Indeed, we have

$$\begin{aligned}\left\langle\frac{\partial\hat{K}_{\lambda 0}}{\partial\beta}\right\rangle &= \left(\frac{\partial\hat{K}_\lambda}{\partial\beta}\right)_0 + \left\langle\sum_{s=1}^{M}\left(\frac{\partial\hat{K}_\lambda}{\partial A_s}\frac{\partial A_{s0}}{\partial\beta} + \frac{\partial\hat{K}_\lambda}{\partial\lambda_s}\frac{\partial\lambda_{s0}}{\partial\beta}\right)\right\rangle_0\\ &= \left(\frac{\partial\hat{H}_\lambda}{\partial\beta}\right)_0 + \sum_{s=1}^{M}\left(\left\langle\frac{\partial\hat{H}_\lambda}{\partial A_s}\right\rangle_0\frac{\partial A_{s0}}{\partial\beta} + \left\langle\frac{\partial\hat{H}_\lambda}{\partial\lambda_s}\right\rangle_0\frac{\partial\lambda_{s0}}{\partial\beta}\right).\end{aligned} \tag{4.50}$$

If $(\partial\hat{K}_\lambda/\partial\beta)_0 = (\partial\hat{H}_\lambda/\partial\beta)_0 = 0$, the above expression vanishes due to Eqs. (4.30) and (4.31), equivalent to Eqs. (4.34). In such a situation, the y-dependence of $\vec{A}_0$ and $\vec{\lambda}_0$ does not spoil the standard thermodynamic relations. Consequently, there is a complete equivalence between the MF statistical mechanics and MF thermodynamics at the level of the first derivatives of the grand potential Ω. However, it is sometimes convenient to work with $\hat{H}_\lambda$, which depends explicitly on β. Then, if the β-dependent part of $\hat{H}_\lambda$ is proportional to the unit matrix, it constitutes additional contribution to total entropy (apart from the single-particle entropy given by S_{vN} (4.4) with $\hat{\rho} = \hat{\rho}_\lambda$). In certain situations, such modification may be required for a more realistic description of the system.

4.6.3 Second derivatives of grand potential

Second derivatives of Ω require greater care. Let $x_1, x_2 \in \{T, V, \mu, \vec{h}\}$. We may use (4.48) to compute $\partial\Omega/\partial x_1$, i.e.,

$$\frac{\partial\Omega}{\partial x_1} = \left(\frac{\partial\mathcal{F}}{\partial x_1}\right)_0. \tag{4.51}$$

Next, we take the derivative with respect to x_2, and finally put $\vec{A} = \vec{A}_0$, $\vec{\lambda} = \vec{\lambda}_0$ again. This yields

$$\begin{aligned}\frac{\partial^2\Omega}{\partial x_2 \partial x_1} &= \frac{\partial}{\partial x_2}\left(\frac{\partial \mathcal{F}}{\partial x_1}\right)_0 = \left(\frac{\partial^2 \mathcal{F}}{\partial x_2 \partial x_1}\right)_0 + \\ &+ \sum_s \left\{ \left(\frac{\partial^2 \mathcal{F}}{\partial x_1 \partial A_s}\right)_0 \frac{\partial A_{s0}}{\partial x_2} + \left(\frac{\partial^2 \mathcal{F}}{\partial x_1 \partial \lambda_s}\right)_0 \frac{\partial \lambda_{s0}}{\partial x_2} \right\}. \end{aligned} \tag{4.52}$$

The above formula is apparently asymmetric with respect to x_1 and x_2. However, it could be shown (cf. Appendix B, i.e., Subsection 11.2) that (4.52) is equivalent to the following expression

$$\begin{aligned}\frac{\partial^2\Omega}{\partial x_2 \partial x_1} &= \frac{\partial}{\partial x_2}\left[\left(\frac{\partial \mathcal{F}}{\partial x_1}\right)_0 + \sum_s \left\{ \left(\frac{\partial \mathcal{F}}{\partial A_s}\right)_0 \frac{\partial A_{s0}}{\partial x_1} + \left(\frac{\partial \mathcal{F}}{\partial \lambda_s}\right)_0 \frac{\partial \lambda_{s0}}{\partial x_1} \right\}\right] = \\ &= \left(\frac{\partial^2 \mathcal{F}}{\partial x_1 \partial x_2}\right)_0 + \sum_s \left\{ \left(\frac{\partial^2 \mathcal{F}}{\partial x_1 \partial A_s}\right)_0 \frac{\partial A_{s0}}{\partial x_2} + \left(\frac{\partial^2 \mathcal{F}}{\partial x_1 \partial \lambda_s}\right)_0 \frac{\partial \lambda_{s0}}{\partial x_2} \right\} \\ &+ \sum_s \left\{ \left(\frac{\partial^2 \mathcal{F}}{\partial x_2 \partial A_s}\right)_0 \frac{\partial A_{s0}}{\partial x_1} + \left(\frac{\partial^2 \mathcal{F}}{\partial x_2 \partial \lambda_s}\right)_0 \frac{\partial \lambda_{s0}}{\partial x_1} \right\} \\ &+ \sum_{s,t} \left\{ \left(\frac{\partial^2 \mathcal{F}}{\partial A_t \partial A_s}\right)_0 \frac{\partial A_{s0}}{\partial x_1}\frac{\partial A_{t0}}{\partial x_2} + \left(\frac{\partial^2 \mathcal{F}}{\partial A_s \partial \lambda_t}\right)_0 \frac{\partial A_{s0}}{\partial x_1}\frac{\partial \lambda_{t0}}{\partial x_2} \right. \\ &+ \left. \left(\frac{\partial^2 \mathcal{F}}{\partial A_s \partial \lambda_t}\right)_0 \frac{\partial A_{s0}}{\partial x_2}\frac{\partial \lambda_{t0}}{\partial x_1} + \left(\frac{\partial^2 \mathcal{F}}{\partial \lambda_s \partial \lambda_t}\right)_0 \frac{\partial \lambda_{s0}}{\partial x_1}\frac{\partial \lambda_{t0}}{\partial x_2} \right\} \\ &+ \sum_s \left\{ \left(\frac{\partial \mathcal{F}}{\partial A_s}\right)_0 \frac{\partial^2 A_{s0}}{\partial x_1 \partial x_2} + \left(\frac{\partial \mathcal{F}}{\partial \lambda_s}\right)_0 \frac{\partial^2 \lambda_{s0}}{\partial x_1 \partial x_2} \right\}. \end{aligned} \tag{4.53}$$

Note, that the last line of the above formula vanishes due to (4.34), but it is explicitly provided for completeness. Eq. (4.53) may be also obtained if we assume that instead of $A_{s0}(x_1, x_2, \ldots)$ and $\lambda_{s0}(x_1, x_2, \ldots)$, we have some arbitrary functions of x_1 and x_2, say $A_{sa}(x_1, x_2, \ldots)$ and $\lambda_{sa}(x_1, x_2, \ldots)$, which are not necessarily the solutions of Eqs. (4.34). Only after computing the second derivative, we eventually put $A_{sa}(\ldots) = A_{s0}(\ldots)$ and $\lambda_{sa}(\ldots) = \lambda_{s0}(\ldots)$.

The first term on the r.h.s. of (4.52) has a structure familiar from standard, non-MF with statistical mechanics. On the other hand, terms in curly brackets are of a purely MF character.

A symmetric form of Eqs. (4.53), together with Eq. (4.48) and (4.52) imply that the standard Maxwell relations, e.g.

$$\begin{aligned}\left(\frac{\partial S}{\partial V}\right)_{T,\mu,\vec{h}} &= \left(\frac{\partial p}{\partial T}\right)_{V,\mu,\vec{h}}, \qquad \left(\frac{\partial N}{\partial V}\right)_{T,\mu,\vec{h}} = \left(\frac{\partial p}{\partial \mu}\right)_{T,V,\vec{h}} \\ \left(\frac{\partial N}{\partial T}\right)_{V,\mu,\vec{h}} &= \left(\frac{\partial S}{\partial \mu}\right)_{T,V,\vec{h}}, \qquad \left(\frac{\partial \vec{M}}{\partial V}\right)_{T,\mu,\vec{h}} = \left(\frac{\partial p}{\partial \vec{h}}\right)_{T,V,\mu} \\ \left(\frac{\partial \vec{M}}{\partial T}\right)_{V,\mu,\vec{h}} &= \left(\frac{\partial S}{\partial \vec{h}}\right)_{T,V,\mu}, \qquad \left(\frac{\partial \vec{M}}{\partial \mu}\right)_{T,V,\vec{h}} = \left(\frac{\partial N}{\partial \vec{h}}\right)_{T,V,\mu}, \end{aligned} \tag{4.54}$$

are valid also in the present case.

4.6.4 Specific heat

For a constant volume and chemical potential, specific heat is defined by

$$\frac{C_{V\mu}}{T} = \left(\frac{\partial S}{\partial T}\right)_{V\mu} = -\left(\frac{\partial^2 \Omega}{\partial T^2}\right)_{V\mu}. \tag{4.55}$$

Using an 'asymmetric' Eq. (4.52) with $x_1 = x_2 = T$, we obtain

$$\left(\frac{\partial S}{\partial T}\right)_{V\mu} = -\left(\frac{\partial^2 \mathcal{F}}{\partial T^2}\right)_0 + \sum_s \left\{ \left(\frac{\partial^2 \mathcal{F}}{\partial T \partial A_s}\right)_0 \frac{\partial A_{s0}}{\partial T} + \left(\frac{\partial^2 \mathcal{F}}{\partial T \partial \lambda_s}\right)_0 \frac{\partial \lambda_{s0}}{\partial T} \right\}. \tag{4.56}$$

First term on the r.h.s of (4.56), i.e., $-(\partial^2 \mathcal{F}/\partial T^2)_0$ corresponds to the expression for C_V/T encountered in a standard statistical mechanics. In the MF case, we have

$$\left(\frac{\partial^2 \mathcal{F}}{\partial T^2}\right)_0 \sim \langle \hat{K}_{\lambda 0}^2 \rangle - \langle \hat{K}_{\lambda 0} \rangle^2 \tag{4.57}$$

This part of (4.56) is related to a change of the population of *fixed* MF energy levels, i.e., eigenstates of $\hat{K}_{\lambda 0}$. The second term on the r.h.s of (4.56) is a genuine MF contribution to specific heat. It is related to the change of the equilibrium MF Hamiltonian and its eigenvalues with the varying temperature. This term is present both in the original Landau formulation [148], where it is responsible for the discontinuity of a specific heat at the phase transition point, as well as in various MF microscopic models. The notable examples include Bragg - Williams approximation [152, 184], or the BCS theory [41, 42, 43, 44].

Note, that $\hat{K}_\lambda$ and its derivatives with respect to A_s and λ_s usually do non commute. Therefore, if we want to express the second part of (4.56) in analogy to (4.49), i.e., by invoking explicitly averages of the relevant operators, the resulting formula, which requires multiple application of Eq. (4.35), becomes quite complicated, and therefore will not be presented here. However, such form is not necessary either, as long as we have an analytical formula for $\mathcal{F}$ at our disposal.

4.6.5 Other thermodynamic potentials

Having defined grand potential Ω (Eq. 4.47), we may obtain other thermodynamic potentials in a standard manner. Helmholtz free energy F is defined as

$$F(T, V, N, \vec{h}) = \Omega\big(T, V, \mu(T, V, N, \vec{h})\big) + \mu(T, V, N, \vec{h}) N. \tag{4.58}$$

Note, that in the above formula and in what follows, by N we denote equilibrium value of the particle number, i.e., $N = A_{s0}$ for $s = 1$. By computing the derivatives of (4.58) with respect to all thermodynamic variables, i.e. T, V, N, $\vec{h}$, etc., it may be verified that the above, standard definition of the free energy is indeed consistent. Namely, having in mind that for equilibrium situation we have $N = -(\partial \mathcal{F}/\partial \mu)_0$, and Eqs. (4.34) are fulfilled (in particular, $0 = (\partial \mathcal{F}/\partial N)_0$), we obtain

$$\begin{aligned} \frac{\partial F}{\partial N} &= +\left(\frac{\partial \mathcal{F}}{\partial \mu}\right)_0 \left(\frac{\partial \mu}{\partial N}\right)_0 + \left(\frac{\partial \mu}{\partial N}\right)_0 N + \mu \\ &+ \sum_s \left\{ \left(\frac{\partial \mathcal{F}}{\partial A_s}\right)_0 \frac{\partial A_{s0}^{(F)}}{\partial N} + \left(\frac{\partial \mathcal{F}}{\partial \lambda_s}\right)_0 \frac{\partial \lambda_{s0}^{(F)}}{\partial N} \right\} = \mu. \end{aligned} \tag{4.59}$$

Superscript (F) indicates that after the Legendre transform $\mu \leftrightarrow N$ is made, the equilibrium values of mean-fields $A_{s0}^{(F)}$ and Lagrange multipliers $\lambda_{s0}^{(F)}$ depend on $T, V, N, \vec{h}$, and not on $T, V, \mu, \vec{h}$. Derivatives with respect to $y = T, V, \vec{h}$ may be computed analogously to $\partial F/\partial N$ Eq. (4.59). In this manner we obtain formulas for the entropy, pressure and magnetization, respectively. The latter quantities can be also defined as average values of the corresponding operators. Both ways are equivalent; this follows from the analogous relations for Ω. Moreover, Eq. (4.59) shows, that indeed μ is the proper chemical potential. This issue is also discussed in. Sec 4.7.

If in Eqs. (4.52) and (4.53) we replace Ω by F, such obtained expressions (now with $x_1, x_2 \in \{T, V, N, \vec{h}\}$) will still be equivalent. This is because μ does depend neither on the components of $\vec{A}$ (apart from $N = A_{1(0)}$), nor on $\vec{\lambda}$. When N, and not μ is an independent variable, the latter is obtained, together with $A^{(F)}_{2(0)}, \ldots, A^{(F)}_{M(0)}$ and $\vec{\lambda}^{(F)}_{s(0)}$, i.e. the remaining components of $\vec{A}_0$ and $\vec{\lambda}_0$, from Eqs. (4.34). Specific heat is now given by

$$\begin{aligned} \frac{C_{VN}}{T} &= \left(\frac{\partial S}{\partial T}\right)_{VN} = -\left(\frac{\partial^2 \mathcal{F}}{\partial T^2}\right)_0 + \left(\frac{\partial^2 \mathcal{F}}{\partial T \partial \mu}\right)_0 \frac{\partial \mu}{\partial T} \\ &+ \sum_s \left\{ \left(\frac{\partial^2 \mathcal{F}}{\partial T \partial A_s}\right)_0 \frac{\partial A^{(F)}_{s0}}{\partial T} + \left(\frac{\partial^2 \mathcal{F}}{\partial T \partial \lambda_s}\right)_0 \frac{\partial \lambda^{(F)}_{s0}}{\partial T} \right\}. \end{aligned} \tag{4.60}$$

Because in the present situation N and T are independent variables, for $s = 1$ we have $\partial A^{(F)}_{1(0)}/\partial T = \partial N/\partial T = 0$. As in a non-MF statistical mechanics, F is more convenient then Ω, because we usually work with fixed N, and not μ. Moreover, working with the lattice models, it is convenient to define a particle density as $n \equiv N/\Lambda$ (where Λ is a number of lattice sites), and not as $\varrho \equiv N/V$. For constant V (rigid lattice), we may put the volume of the unit cell equal unity, then $V = \Lambda$ and $\varrho = n$. However, if we want to study thermal expansion of the lattice, it is much more convenient to work with variable V and the constant pressure and particle number. In complete analogy to the standard thermodynamics, we can make another Legendre transform and obtain Gibbs potential (free enthalpy) G as

$$\begin{aligned} G(T, p, N, \vec{h}) &= F(T, V, N, \vec{h}) + pV = \Omega\big(T, V(T, p, N, \vec{h}), \mu(T, p, N, \vec{h}), \vec{h}\big) \\ &+ \mu(T, p, N, \vec{h}) N + pV(T, p, N, \vec{h}). \end{aligned} \tag{4.61}$$

In such situation, T, p, N and $\vec{h}$ are the independent thermodynamic variables. Again, one can easily check that this definition is consistent, e.g. all standard relations for the first derivatives of G remain valid.

4.6.6 Thermodynamic equilibrium: final remarks

It should be noted, that application of MF thermodynamics requires always some care. Namely, certain relations, encountered in standard thermodynamics and statistical mechanics, may cease to be valid in the MF case. For example, it is well-known that the following conditions,

$$\left(\frac{\partial p}{\partial V}\right)_T < 0, \qquad C_p > C_V > 0. \tag{4.62}$$

expressing the stability of matter, should be fulfilled. This is indeed always the case for a non-MF Hamiltonians. However, for MF models, due to presence of additional terms in the second derivatives of thermodynamic potentials (cf. the second line of Eq. (4.52)), at least one of the above inequalities may be violated. Similar situation is present e.g. in the famous van der Walls description of the hard-sphere gas, or Bragg-Williams model of binary alloy; in both cases we must invoke Maxwell construction to cure the unphysical predictions of the model (cf. [146]).

Finally, let us note, that the mutual consistency of statistical mechanics and thermodynamics within the MF description has been analyzed also in Ref. [185], nonetheless, from a different perspective then presented here.

4.7 Additional remarks on chemical potential

Within the grand canonical ensemble, a number of particles is fluctuating, but its average value is fixed and *a priori* given. In analogy to the case of non-mean-field formalism, also in the present situation we have achieved this by introducing the chemical potential μ. On the other hand, MF Hamiltonians frequently depend on the average particle number, i.e., the latter is one of the relevant mean-fields (denoted A_1), and is treated as a variational parameter within our method. Consequently, the corresponding Lagrange multiplier $\lambda_1 \equiv \lambda_N$ must be introduced. Hence, in order to obtain the solution corresponding to the equilibrium situation, $\mathcal{F}$ (which depends on μ, A_1, and λ_N) has to be minimized also with respect to A_1. Here we provide some additional arguments for the correctness of such formulation.

Among $2M$ equations[33] (4.34), we have to solve the following

$$\left(\frac{\partial \mathcal{F}}{\partial N}\right)_0 = 0, \qquad \left(\frac{\partial \mathcal{F}}{\partial \lambda_N}\right)_0 = 0 \Leftrightarrow \mathrm{Tr}[\hat{N}\hat{\rho}_\lambda]_0 = A_{1(0)}(\mu) = N(\mu). \tag{4.63}$$

Solution of (4.63) yields $N(\mu)$, which corresponds to the minimum of $\mathcal{F}$ and which at the same time, according to Eqs. (4.8), is the average of the corresponding operator. μ is determined from the condition, that $N(\mu)$ should be equal to the *a priori* given value.

As discussed above, μ is true chemical potential, i.e., it appears in thermodynamic relations. Explicitly, invoking (4.48) for $y = \mu$, we have

$$\frac{\partial \Omega}{\partial \mu} = \left(\frac{\partial \mathcal{F}}{\partial \mu}\right)_0 + \left(\frac{\partial \mathcal{F}}{\partial \vec{A}}\right)_0 \dot{\vec{A}}_0 + \left(\frac{\partial \mathcal{F}}{\partial \vec{\lambda}}\right)_0 \dot{\vec{\lambda}}_0 = \left(\frac{\partial \mathcal{F}}{\partial \mu}\right)_0 = -N, \tag{4.64}$$

where $\dot{\vec{A}}_0 \equiv \partial \vec{A}_0/\partial\mu$ and $\dot{\vec{\lambda}}_0 \equiv \partial \vec{\lambda}_0/\partial\mu$. Similarly, for $F = \Omega + \mu N$, treated as a function of its natural variable N, it can easily be shown that $\partial F/\partial N = \mu$, cf. Eq. (4.59). Neither λ_N, nor $\tilde{\mu} \equiv \mu + \lambda_N$ have this property. Note also, that among $2M$ equations (4.34), only one, i.e.,

$$\left(\frac{\partial \mathcal{F}}{\partial A_1}\right)_0 = \left\langle \frac{\partial \hat{H}}{\partial A_1} \right\rangle + \lambda_N = 0, \tag{4.65}$$

contains λ_N not in the combination $\mu + \lambda_N = \tilde{\mu}$. After change of variables $(\mu, \lambda_N) \to (\tilde{\mu}, \lambda_N)$, the remaining $2M - 1$ equations which contain only $\tilde{\mu}$ may be solved separately, and the value of λ_N (and, consequently, also μ) may be found using (4.65).

Alternatively, if A_1 was not treated as a variational parameter, in order to ensure self-consistency it is sufficient to introduce only one Lagrange multiplier $\tilde{\mu}$ related to A_1. Then we deal with the reduced problem involving $2M - 1$ equations (4.34), but without Eq. (4.65). Consequently, the resulting equilibrium values of the remaining mean fields $A_{2(0)}, \ldots, A_{M(0)}$ and $\vec{\lambda}_0$ are the same as in the former case. However, the latter way is less satisfactory, as without making use of (4.65) we obtain additional terms in the standard thermodynamic relations, (e.g. Eq. (4.64) is no longer valid).

Now, let us consider briefly non-equilibrium situation. Although for A_1 the thermodynamic fluctuations should be usually insignificant, such analysis may also clarify the relation between μ and λ_N. For a given value of A_1 and μ we solve (4.41), this yields $\lambda_N = \lambda_N(\vec{A})$. The latter condition, however, obviously does not determine the value of μ, as we have one equation

$$N = \frac{\mathrm{Tr}[\hat{N} \exp\big(-\beta(\hat{H}_\lambda - (\mu + \lambda_N)\hat{N}\big)]}{\mathrm{Tr}[\exp\big(-\beta(\hat{H}_\lambda - (\mu + \lambda_N)\hat{N}\big)]}, \tag{4.66}$$

[33] Once again, for simplicity, we assume $b_1 = \ldots = b_P = 0$.

from which we can determine only $\mu + \lambda_N = \tilde{\mu}$. On the other hand, even if the factor $\exp(-\beta\lambda_N A_1)$ cancels out in $\hat{\rho}_z(\vec{A})$, and does not appear in (4.66), λ_N is necessary to obtain the value of $\mathcal{F}$. The natural solution of this problem is to use value of μ given by (4.34) solved for μ, and not for N. Then, the chemical potential becomes $\vec{A}$-independent quantity.

Note, however, that it is $\tilde{\mu}$, not μ, which appears as a pre-factor multiplying $\hat{N}$ in $\hat{K}_\lambda$, and consequently, in Eq. (4.66).[34] In some situations the symmetry arguments may allow to determine value of $\tilde{\mu}$, but not μ. For example, in the case of a single-band lattice-fermion models with symmetric dispersion relation and at half-filling (one electron per site), it follows that $\tilde{\mu} = 0$, unlike the standard situation when $\mu = 0$.

The arguments similar to those given above would remain valid in case of any other observable, which average value is *a priori* known. On the technical level, this prior information may be taken into account by using a chemical potential -like term. In particular, the above discussion may be easily generalized to the situation with more then one kind of particles is present.

4.8 Equivalence classes of mean-field Hamiltonians

In Subsection 4.5 we have introduced self-consistent MF Hamiltonian $\hat{H}_z(\vec{A})$ (4.42). It turns out, that analysis of the properties of both $\hat{H}_z(\vec{A})$ and the related density operator $\hat{\rho}_z(\vec{A})$ (4.43) provides deeper insight into the structure of the present formalism.[35] Let us now consider a MF Hamiltonian $\hat{H}(\vec{A})$ (4.1) of the following form

$$\hat{H}(\vec{A}) = \hat{H}_{e0} + C_0(\vec{A}) \cdot \hat{\mathbf{1}}_D + \sum_{s=1}^{M} C_s(\vec{A})\hat{A}_s + \sum_{w=1}^{M'} G_w(\vec{A})\hat{B}_w. \tag{4.67}$$

In the above, $\hat{\mathbf{1}}_D$ is an unit operator, $\hat{H}_{e0}$ is an $\vec{A}$-independent part, whereas $C_0(\vec{A})$, $C_s(\vec{A})$ and $G_w(\vec{A})$ are complex-valued functions of mean-fields $\vec{A}$. We make a distinction between $\hat{A}_s$ and $\hat{B}_w$ operators, $\hat{H}(\vec{A})$ does not depend on the average values of the latter, i.e., $\forall w,\ B_w \equiv \langle \hat{B}_w \rangle \notin \{A_1, \ldots, A_M\}$. From (4.15) and (4.16), for $\hat{H}(\vec{A})$ (4.67) we obtain

$$\begin{aligned} \hat{K}_\lambda(\vec{A}, \vec{\lambda}) &= \hat{H}(\vec{A}) - \sum_{s=1}^{M} \lambda_s(\hat{A}_s - A_s) - \mu\hat{N} \\ &= \hat{K}_{e0} + C_0^\lambda(\vec{A}, \vec{\lambda}) \cdot \hat{\mathbf{1}}_D + \sum_{s=1}^{M} C_s^\lambda(\vec{A}, \vec{\lambda})\hat{A}_s + \sum_{w=1}^{M'} G_w(\vec{A})\hat{B}_w, \end{aligned} \tag{4.68}$$

with $\hat{K}_{e0} = \hat{H}_{e0} - \mu\hat{N}$, and

$$C_0^\lambda(\vec{A}, \vec{\lambda}) = C_0(\vec{A}) + \sum_{s=1}^{M} \lambda_s A_s, \quad C_s^\lambda(\vec{A}, \vec{\lambda}) = C_s(\vec{A}) - \lambda_s. \tag{4.69}$$

For a given $\vec{A}$, the self-consistency equations (4.41) may be rewritten as

$$\begin{aligned} A_t &= \mathrm{Tr}[\hat{A}_t\hat{\rho}_\lambda(\vec{A}, \vec{\lambda})] = \langle \hat{A}_t \rangle \\ &= \frac{\mathrm{Tr}[\hat{A}_t \exp\big(-\beta(\hat{K}_{e0} + \sum_{s=1}^{M} C_s^\lambda(\vec{A})\hat{A}_s + \sum_{w=1}^{M'} G_w(\vec{A})\hat{B}_w)\big)]}{\mathrm{Tr}[\exp\big(-\beta(\hat{K}_{e0} + \sum_{s=1}^{M} C_s^\lambda(\vec{A})\hat{A}_s + \sum_{w=1}^{M'} G_w(\vec{A})\hat{B}_w)\big)]}. \end{aligned} \tag{4.70}$$

[34] This is true as long as the MF Hamiltonian $\hat{H}(\vec{A})$ does not contain a term $\sim \hat{N}$. On the other hand, if such term (say $C_1(\vec{A})\hat{N}$) is present, then in the corresponding $\hat{K}_\lambda$, $\hat{N}$ is multiplied by $C_1(\vec{A}) + \tilde{\mu}$.

[35] For simplicity, again, in the large part of the following discussion, we ignore variational parameters $\vec{b}$ of a non-MF character. However, modifications caused by the presence of such variables are briefly analyzed in Sub-subsection 4.8.8.

for $t = 1, 2, \ldots, M$. Note also, that the $C_0^\lambda(\vec{A}, \vec{\lambda})\hat{\mathbf{1}}_D$ term cancels out. When Eqs. (4.70) are solved[36] for $\vec{\lambda}$, we have $\vec{\lambda} = \vec{\lambda}(\vec{A})$, and hence we obtain $C_s^\lambda(\vec{A}, \vec{\lambda}(\vec{A})) \equiv \mathcal{C}_s^\lambda(\vec{A})$. In other words, in the $2M$-dimensional space of the mean-field variables $(\vec{A}, \vec{\lambda})$, we remain on the hyper-surface defined by Eqs. (4.41), with $\vec{A} \in \mathcal{D}_\mathcal{A}$.

4.8.1 Universality classes of mean-field Hamiltonians

Let us change the original form of the $C_s(\vec{A})$ coefficients, according to

$$C_s(\vec{A}) \to \tilde{C}_s(\vec{A}) \qquad s = 0, 1, \ldots, M, \tag{4.71}$$

but without modifying any of the M' functions $G_w(\vec{A})$. Such step leads to new MF Hamiltonian, $\hat{H}^{(\sim)}(\vec{A})$, for which the corresponding Hamiltonian (4.15), denoted here as $\hat{H}_\lambda^{(\sim)}(\vec{A}, \vec{\lambda})$, may be constructed according to (4.68). Next, we may rewrite Eqs. (4.70) with this new $\hat{H}_\lambda^{(\sim)}(\vec{A}, \vec{\lambda})$, and solve for $\lambda_1, \ldots, \lambda_M$, which yields $\lambda_1 = \tilde{\lambda}_1(\vec{A})$, $\ldots$, $\lambda_M = \tilde{\lambda}_M(\vec{A})$. Note, that in order to fulfill Eqs. (4.70), numerical values of the C_s^λ coefficients must remain unchanged by transformations (4.71), i.e., for $s = 1, \ldots, M$ we must have

$$\mathcal{C}_s^\lambda(\vec{A}) \equiv C_s(\vec{A}) - \lambda_s(\vec{A}) = \tilde{C}_s(\vec{A}) - \tilde{\lambda}_s(\vec{A}) \equiv \tilde{\mathcal{C}}_s^\lambda(\vec{A}). \tag{4.72}$$

In result, the $\vec{A}$-dependence of all $\mathcal{C}_s^\lambda(\vec{A})$ coefficients is completely determined by the choice of operators $\hat{H}_{e0}$, $\{\hat{A}_s\}_{s=1}^M$, and $\{\hat{B}_w\}_{w=1}^{M'}$, as well as the functional form of the $G_w(\vec{A})$ functions. The only $\vec{A}$-dependent part of such obtained $\hat{H}_z(\vec{A}) \equiv \hat{H}_\lambda(\vec{A}, \vec{\lambda}(\vec{A}))$ (4.42), that may be non-trivially modified by (4.71), is the $C_0^\lambda(\vec{A}, \vec{\lambda}(\vec{A}))\cdot\hat{\mathbf{1}}_D \equiv \mathcal{C}_0^\lambda(\vec{A})\cdot\hat{\mathbf{1}}_D$ term. Therefore, $\mathcal{F}(\vec{A}, \vec{\lambda}(\vec{A})) \equiv \mathcal{F}_z(\vec{A})$ (4.44) is usually also altered by transformations (4.71). However, density operator $\hat{\rho}_z(\vec{A})$ (4.43) is invariant under (4.71). Consequently, for a given $\vec{A}$, all the averages, also the averages of the $\hat{H}_{e0}$ and $\hat{B}_w$ operators are invariant, too.

We will say that all $\hat{H}(\vec{A})$ MF Hamiltonians (4.67) (hence also all the corresponding $\hat{H}_\lambda(\vec{A}, \vec{\lambda})$ and $\hat{H}_z(\vec{A})$ Hamiltonians), which are constructed from the same set of operators ($\hat{H}_{e0}$, $\{\hat{A}_s\}_{s=1}^M$, and $\{\hat{B}_w\}_{w=1}^{M'}$), and by using the same set of $G_w(\vec{A})$ functions, belong to the same *universality class*.[37] MF Hamiltonians $\hat{H}_z(\vec{A})$ from the same universality class may differ only by the term proportional to unit an operator.

4.8.2 Equivalence relation

Let us consider now two MF Hamiltonians, $\hat{H}(\vec{A})$ and $\hat{H}^{(\sim)}(\vec{A})$, which both belong to the same universality class. In the previous Sub-subsection we have shown, that after Eqs. (4.70) are solved, the corresponding self-consistent Hamiltonians $\hat{H}_z(\vec{A})$ and $\hat{H}_z^{(\sim)}(\vec{A})$ may differ only by a $\mathcal{C}_0^\lambda(\vec{A}) = C_0^\lambda(\vec{A}, \vec{\lambda}(\vec{A}))$ term. However, we have also

$$\begin{aligned} \mathcal{C}_0^\lambda(\vec{A}) \;&\equiv\; C_0(\vec{A}) + \sum_{s=1}^M \lambda_s(\vec{A})A_s = \sum_{s=1}^M \big(\lambda_s(\vec{A})A_s - C_s(\vec{A})A_s\big) \\ &+\; C_0(\vec{A}) + \sum_{s=1}^M C_s(\vec{A})A_s = C_0(\vec{A}) + \sum_{s=1}^M C_s(\vec{A})A_s - \sum_{s=1}^M \mathcal{C}_s^\lambda(\vec{A})A_s. \end{aligned} \tag{4.73}$$

[36] We assume that the solution is unique, or if there exist a finite number of solutions, and one of them may be unambiguously selected, cf. Subsection 4.5.

[37] This term should not be confused with the one appearing in the renormalization group theory, cf. e.g. [132].

Because for $s > 0$, the coefficients $\mathcal{C}_s^\lambda(\vec{A})$ are invariants of transformations (4.71), cf. Eq. (4.72), it follows that $\hat{H}_z(\vec{A}) = \hat{H}_z^{(\sim)}(\vec{A})$ if only

$$C_0(\vec{A}) + \sum_{s=1}^{M} C_s(\vec{A}) A_s = \tilde{C}_0(\vec{A}) + \sum_{s=1}^{M} \tilde{C}_s(\vec{A}) A_s. \tag{4.74}$$

As a consequence of (4.74), we have also

$$U(\vec{A}) \equiv \langle \hat{H}_z(\vec{A}) \rangle_z \equiv \mathrm{Tr}[\hat{\rho}_z(\vec{A}) \cdot \hat{H}_z(\vec{A})] = \tilde{U}(\vec{A}) \tag{4.75}$$

as well as

$$\mathcal{F}_z(\vec{A}) \equiv -\beta^{-1} \ln(\mathrm{Tr}[e^{-\beta(\hat{H}_z - \mu\hat{N})}]) = \mathcal{F}_z^{(\sim)}(\vec{A}) \tag{4.76}$$

Note, that μ is treated as independent variable, and its value (as well as the value of $\mu\langle\hat{N}\rangle = \mu A_1$ term) is identical for all such transformed grand Hamiltonians. Consequently, equilibrium values of mean fields are invariants of those transformations (4.71), for which the condition (4.74) is obeyed. In other words, all MF Hamiltonians constructed from the same set of operators $\hat{H}_{e0}$, $\{\hat{A}_s\}_{s=1}^M$ and $\{\hat{B}_w\}_{w=1}^{M'}$ by using the same set of $G_w(\vec{A})$ functions, and additionally characterized by the same $\vec{A}$-dependence of $\langle \hat{H}_z(\vec{A}) \rangle_z = U(\vec{A})$, are equivalent. In this manner we have identified the equivalence relation, which divides MF Hamiltonian into *equivalence classes*. Transformations (4.71) may be viewed in analogy to gauge transformations, not changing the physical contents of the model.

Let us also note at this point, that for different Hamiltonians from the same equivalence class, we would obtain, in general, different Bogoliubov-de Gennes equations (4.38), leading to different optimal values of mean-fields ($\vec{A}_{sc}^{(0)}$). This is another serious drawback of the BdG non-variational method.

4.8.3 Special case of transformations (4.71)

For our purposes, it is sufficient to consider a special class of transformations (4.71), given by

$$\begin{aligned} \tilde{C}_s(\vec{A}) &= \kappa_s C_s(\vec{A}), \qquad \tilde{C}_0(\vec{A}) = C_0(\vec{A}) + (1-\kappa_s) C_s(\vec{A}) A_s, \\ \tilde{C}_t(\vec{A}) &= C_t(\vec{A}) \quad \text{for} \quad t = 1, \ldots, M, \quad t \neq s, \end{aligned} \tag{4.77}$$

and denoted by $\mathcal{R}_s(\kappa_s)$. For each s, transformations $\{\mathcal{R}_s(\kappa_s)\}, s = 1, \ldots, M$ form an Abelian group, $\mathcal{G}_s$. Indeed, it may be easily shown that $\mathcal{R}_s(\kappa_s)\mathcal{R}_s(\omega_s) = \mathcal{R}_s(\kappa_s\omega_s)$, the unity of the group is $e_s = \mathcal{R}_s(1)$ and that $\mathcal{R}_s^{-1}(\omega_s) = \mathcal{R}_s(\omega_s^{-1})$. The complete group of transformation is a direct product of M groups $\mathcal{G}_s$, $s = 1, 2, \ldots, M$. Obviously, for each $\mathcal{R}_s(\kappa_s)$, as well as for the composition of arbitrary number of such transformations, the condition (4.74) is fulfilled, therefore also $U(\vec{A}) = \langle \hat{H}_z(\vec{A}) \rangle$ remains unchanged.

4.8.4 Reduced form of mean-field Hamiltonian

Let us consider now a MF Hamiltonian (4.67), for which

$$\hat{H}_{e0} = 0, \qquad \text{and} \qquad \forall_w : \qquad G_w = 0. \tag{4.78}$$

We can apply a sequence of M transformations $\mathcal{R}_s(\kappa_s)$ (4.77), $s = 1, \ldots, M$, with $\kappa_s = 0$, i.e.,

$$C_s(\vec{A}) \to \tilde{C}_s(\vec{A}) = 0, \quad s \neq 0, \qquad C_0(\vec{A}) \to \tilde{C}_0(\vec{A}) = C_0(\vec{A}) + \sum_s C_s(\vec{A}) A_s. \tag{4.79}$$

This yields $\hat{H}^{(\sim)}(\vec{A})$ proportional to the unit matrix, $\hat{H}^{(\sim)}(\vec{A}) = W(\vec{A})\hat{\mathbf{1}}_D$, where

$$W(\vec{A}) = \langle \hat{H}(\vec{A}) \rangle = \tilde{C}_0(\vec{A}). \tag{4.80}$$

Consequently, we obtain $\hat{K}_\lambda$ (4.16) of the following form

$$\hat{K}_\lambda^{(\sim)}(\vec{A}, \vec{\lambda}) = \hat{H}_\lambda^{(\sim)}(\vec{A}, \vec{\lambda}) - \mu\hat{N} = -\sum_{s=1}^{M} \lambda_s(\hat{A}_s - A_s) + W(\vec{A})\hat{\mathbf{1}}_D - \mu\hat{N}. \tag{4.81}$$

Next, using $\hat{K}_\lambda^{(\sim)}(\vec{A}, \vec{\lambda})$, the corresponding density operator $\hat{\rho}_\lambda(\vec{A}, \vec{\lambda}) = \hat{\rho}_\lambda^{(\sim)}(\vec{A}, \vec{\lambda})$ and partition function $\mathcal{Z}_\lambda(\vec{A}, \vec{\lambda}) = \mathcal{Z}_\lambda^{(\sim)}(\vec{A}, \vec{\lambda})$ (4.29), as well as generalized grand potential $\mathcal{F}(\vec{A}, \vec{\lambda})$ (4.33) may be constructed. From the previous discussion it follows, that if the conditions (4.41) are fulfilled, then $\hat{K}_\lambda$ (4.68) for which the conditions (4.78) are fulfilled, is completely equivalent to (4.81). Namely, those two Hamiltonians lead to identical $\hat{H}_z(\vec{A})$, $\hat{\rho}_z(\vec{A})$ and $\mathcal{F}_z(\vec{A})$, therefore they also yield the same optimal values of mean fields $\vec{A}_0$ and $\vec{\lambda}_0 = \vec{\lambda}(\vec{A}_0)$ ($\vec{A}_0$ may be formally obtained from the minimization of $\mathcal{F}_z(\vec{A})$), and the same equilibrium MF (grand) Hamiltonians $\hat{K}_{\lambda 0}$.

Above, we have formally determined the optimal values of the mean-fields using a two-step procedure. First, for a fixed value of $\vec{A}$, we have eliminated $\lambda_1, \ldots, \lambda_M$ using Eqs. (4.31), which yields $\vec{\lambda} \equiv \vec{\lambda}(\vec{A})$. Then, we have subsequently solved Eqs. (4.30) in order to obtain $\vec{A} = \vec{A}_0$. However, because the complete solution is determined by both (4.30) and (4.31), or, equivalently, by $2M$ equations[38] (4.34), we may have solved first Eqs. (4.30). For the Hamiltonian (4.81), we obtain then

$$\lambda_s = \lambda_s^\star(\vec{A}) = -\frac{\partial W(\vec{A})}{\partial A_s}, \quad s = 1, 2, \ldots, M, \tag{4.82}$$

because $\partial W(\vec{A})/\partial A_s = \partial \langle \hat{H}(\vec{A}) \rangle / \partial A_s = \langle \partial \hat{H}(\vec{A})/\partial A_s \rangle$ if only $C_s(\vec{A}) \neq 0$. In contrast to the case of Eqs. (4.31), the analytic solutions of Eqs. (4.82) (denoted by $\lambda_s^\star(\vec{A})$) may be easily obtained, if only an explicit form of $W(\vec{A})$ is known. Using $\vec{\lambda}^\star(\vec{A}) \equiv (\lambda_1^\star(\vec{A}), \ldots, \lambda_M^\star(\vec{A}))$ and $\hat{H}_\lambda^{(\sim)}(\vec{A}, \vec{\lambda})$ (4.81), we define

$$\hat{H}_\lambda^{(\star)}(\vec{A}) = \hat{H}_\lambda^{(\sim)}(\vec{A}, \vec{\lambda}^\star(\vec{A})) = \sum_{s=1}^{M} \frac{\partial W(\vec{A})}{\partial A_s}(\hat{A}_s - A_s) + W(\vec{A}) \cdot \hat{\mathbf{1}}_D. \tag{4.83}$$

Next, this Hamiltonian may be used to obtain the following density operator

$$\hat{\rho}_\lambda^{(\star)}(\vec{A}) \equiv \hat{\rho}_\lambda(\vec{A}, \vec{\lambda}^\star(\vec{A})), \tag{4.84}$$

for $\hat{\rho}_\lambda(\vec{A}, \vec{\lambda})$ (4.29). Using $\hat{\rho}_\lambda^{(\star)}(\vec{A})$, in turn, we may determine the equilibrium values of $\vec{A}$ by solving (usually by means of numerical analysis) the remaining self-consistent equations (4.8).

Summarizing, we may eliminate $\vec{\lambda}$ in favor of $\vec{A}$ by solving Eqs. (4.30), i.e., Eqs. (4.70). This yields M functions $\lambda_s = \lambda_s(\vec{A})$, $s = 1, \ldots, M$. However, usually an explicit analytic form of $\vec{\lambda}(\vec{A}) \equiv (\lambda_1(\vec{A}), \ldots, \lambda_M(\vec{A}))$ cannot be obtained. Alternatively, we may eliminate $\vec{\lambda}$ using another condition, i.e., Eqs. (4.30), and for MF Hamiltonian of the form (4.81) we obtain $\lambda_s = \lambda_s^\star(\vec{A})$ (4.82). For $\hat{H}(\vec{A})$ of the form (4.78), by using transformations (4.79), we arrive at the equivalent Hamiltonian $\hat{H}^{(\sim)}(\vec{A}) = W(\vec{A})\hat{\mathbf{1}}_D$, and the corresponding Hamiltonian $\hat{H}_\lambda^{(\star)}(\vec{A})$ (4.83), i.e. $\hat{H}_\lambda(\vec{A}, \vec{\lambda})$ (4.81) with $\lambda_s = \lambda_s^\star(\vec{A})$ given by (4.82). Obviously, in general,

[38] For simplicity, $\vec{b}$ variables have been ignored.

$\lambda_s(\vec{A}) \neq \lambda_s^\star(\vec{A})$. Consequently, $\hat{H}_\lambda^{(\star)}(\vec{A})$ (4.83) differs from $\hat{H}_z(\vec{A})$ (4.42). $\hat{H}_\lambda^{(\star)}(\vec{A})$ may be used only in order to obtain the equilibrium values of mean-fields, but has no interpretation for $\vec{A} \neq \vec{A}_0$. Nevertheless, we have

$$\lambda_s(\vec{A}_0) = \lambda_s^\star(\vec{A}_0) = -\left(\frac{\partial W(\vec{A})}{\partial A_s}\right)_0 = \lambda_{s0}. \tag{4.85}$$

Transformations (4.79), leading to Hamiltonian of the form (4.81) are very convenient, as they reduce (by a factor of two) the number of equations to be solved numerically. Also, we have an explicit analytical expression for $W(\vec{A}) = U(\vec{A})$ at our disposal. Therefore, some of the properties of the MF model may be inferred solely from the analysis of functional dependence of $W(\vec{A})$ on mean-fields $\vec{A}$. Moreover, transformations (4.79) allow to compare our formalism with alternative approaches proposed by other Authors.

4.8.5 Present approach and formalism of Reference [79]

In the ($\beta \to \infty$) limit, the present finite-temperature approach becomes equivalent to the true $T = 0$ analysis, and we may compare it directly e.g. with the formalism of Ref. [79]. Hamiltonian of that Reference rewritten in our notation may be recast in the following general form

$$\hat{H}_\lambda^{(K)}(\vec{A}) = \sum_{s=1}^{M} \frac{\partial W(\vec{A})}{\partial A_s}\hat{A}_s = -\sum_{s=1}^{M} \lambda_s^\star(\vec{A})\hat{A}_s. \tag{4.86}$$

$\hat{H}_\lambda^{(K)}(\vec{A})$ differs from $\hat{H}_\lambda^\star(\vec{A})$ (4.83) only by the term proportional to the unit matrix, i.e., $(\sum_{s=1}^{M} \lambda_s^{(\star)}(\vec{A})A_s + \langle\hat{H}(\vec{A})\rangle) \cdot \hat{\mathbf{1}}_D$. Consequently, both (4.83) and (4.86) have the same eigenvectors. Presence of $-\mu\hat{N}$ term does not change this argument, moreover, when applying transformations (4.79) we may either transform this term according to $\mu\hat{N} \to \mu N$, or leave it in the initial form; both ways are equivalent. $W(\vec{A}) \equiv W^{(K)}(\vec{A})$ of Ref. [79] has the following form

$$W^{(K)}(\vec{A}) \equiv \langle\hat{H}_{tJ}\rangle_C^{app} = \langle\Psi_0|\hat{H}_{tJ}^{(ren)}|\Psi_0\rangle \approx \frac{\langle\Psi_0|\hat{P}_C\hat{H}_{tJ}\hat{P}_C|\Psi_0\rangle}{\langle\Psi_0|\hat{P}_C^2|\Psi_0\rangle}, \tag{4.87}$$

with $\hat{H}_{tJ}^{(ren)}$ given by (7.148). The optimal values of the mean-fields ($\vec{A} = \vec{A}_0$) are obtained by solving the self-consistency conditions (4.8). In order to do this, it is legitimate to extend the treatment of Ref. [79] to finite, albeit very low temperatures. The grand-canonical mean-field density operator $\hat{\rho}_\lambda^\star(\vec{A})$ (4.84), constructed either with the help of our $\hat{H}_\lambda^\star(\vec{A})$ (4.83) or $\hat{H}_\lambda^{(K)}(\vec{A})$ (4.86) are identical, because the term proportional to unit matrix cancels out. Therefore, both approaches lead to the same values of $\vec{A}_0$.

Finally, let us note, that in Ref. [79], $\hat{H}_\lambda^{(K)}(\vec{A})$ (4.86) has been derived by varying $W^{(K)}(\vec{A})$ (4.87) with respect to $\langle\Psi_0|$, and by using the relation $\langle\Psi_0|\hat{A}_s|\Psi_0\rangle = A_s$. In that manner, the following Schrödinger equation is obtained,

$$\frac{\partial W^{(K)}(\vec{A})}{\partial\langle\Psi_0|} = \sum_s \frac{\partial W^{(K)}(\vec{A})}{\partial A_s}\frac{\partial A_s}{\partial\langle\Psi_0|} = \sum_s \frac{\partial W^{(K)}(\vec{A})}{\partial A_s} \cdot \hat{A}_s|\Psi_0\rangle, \tag{4.88}$$

in which $\hat{H}_\lambda^{(K)}(\vec{A})$ (4.86) plays the role of an effective MF Hamiltonian.

4.8.6 Generalization of transformations (4.77) to arbitrary form of mean-field Hamiltonian

Apparently, if the conditions (4.78) are not fulfilled, then Eqs. (4.30) are not equivalent to Eqs. (4.82), and we cannot eliminate Lagrange multipliers in an analytic fashion by solving Eqs.

(4.30). However, any $\hat{H}(\vec{A})$ (4.67) may be easily given the form (4.78). In order to do that, we introduce $M'+1$ new mean-fields $B_1, B_2, \ldots, B_{M'}, B_{M'+1}$, $(B_1, B_2, \ldots, B_{M'}, B_{M'+1}) \equiv \vec{B}$. Similarly as in the case of A_s mean-fields, we demand that after the self-consistency conditions are fulfilled, $B_w \equiv \langle \hat{B}_w \rangle$ for $w = 1, \ldots, M'$ and $B_{M'+1} \equiv \langle \hat{H}_{e0} \rangle$, where the averages $\langle \ldots \rangle$ are defined with the help of a density operator, which will be subsequently specified. Therefore, for the newly introduced mean-fields we should repeat the step leading from $\hat{H}$ to $\hat{H}_\lambda$ (4.15), i.e., introduce $M'+1$ Lagrange multipliers $l_1, l_2, \ldots, l_{M'}, l_{M'+1}$ and supplement $\hat{H}(\vec{A})$ (4.67) with the constraint terms (4.13). We obtain

$$\hat{H}_{g\lambda}(\vec{A}, \vec{B}, \vec{\lambda}, \vec{l}) \equiv \hat{H}(\vec{A}) - \sum_{s=1}^{M} \lambda_s (\hat{A}_s - A_s) - \sum_{w=1}^{M'+1} l_w (\hat{B}_w - B_w). \tag{4.89}$$

Now, $M + M' + 1$ Eqs. (4.30) and $M + M' + 1$ Eqs. (4.31) for $\hat{H}_{g\lambda}$ can be written respectively as

$$\begin{aligned} \left\langle \frac{\partial \hat{H}_{g\lambda}(\ldots)}{\partial A_s} \right\rangle &= \left\langle \frac{\partial \hat{H}_{\lambda}(\ldots)}{\partial A_s} \right\rangle = 0, \quad s = 1, \ldots, M; \\ \left\langle \frac{\partial \hat{H}_{g\lambda}(\ldots)}{\partial B_w} \right\rangle &= 0, \quad w = 1, \ldots, M'+1, \end{aligned} \tag{4.90}$$

$$\begin{aligned} \left\langle \frac{\partial \hat{H}_{g\lambda}(\ldots)}{\partial \lambda_s} \right\rangle &= \left\langle \frac{\partial \hat{H}_{\lambda}(\ldots)}{\partial \lambda_s} \right\rangle = \langle \hat{A}_s \rangle - A_s = 0, \quad s = 1, \ldots, M; \\ \left\langle \frac{\partial \hat{H}_{g\lambda}(\ldots)}{\partial l_w} \right\rangle &= \langle \hat{B}_w \rangle - B_w = 0, \quad w = 1, \ldots, M'+1. \end{aligned} \tag{4.91}$$

In the above, all averages $\langle \hat{B}_w \rangle$ are defined with the help of $\hat{\rho}_{g\lambda}(\vec{A}, \vec{B}, \vec{\lambda}, \vec{l})$ given by Eq. (4.29), in which $\hat{H}_\lambda$ is replaced by $\hat{H}_{g\lambda}$ (4.89). Please note, that first M equations of both (4.90) and (4.91) are identical with the corresponding equations for $\hat{H}_\lambda(\vec{A}, \vec{\lambda})$ (4.68). The former may be solved for $\vec{\lambda}$, this yields $\vec{\lambda} \equiv \vec{\lambda}_{vAB}(\vec{A}, \vec{B}) \equiv \vec{\lambda}_{vA}(\vec{A})$. On the other hand, for $w = 1, 2, \ldots, M'+1$, we have

$$\left\langle \frac{\partial \hat{H}_{g\lambda}(\ldots)}{\partial B_w} \right\rangle = l_w + \left\langle \frac{\partial \hat{H}(\vec{A})}{\partial B_w} \right\rangle = l_w = 0. \tag{4.92}$$

Consequently, for each w, we have also $l_{w0} = 0$. Using the latter result, we obtain

$$\hat{H}_{g\lambda}(\vec{A}, \vec{B}, \vec{\lambda}_{AB}(\vec{A}, \vec{B}), \vec{l}_0) = \hat{H}_{g\lambda}(\vec{A}, \vec{B}, \vec{\lambda}_A(\vec{A}), \vec{0}) = \hat{H}_\lambda(\vec{A}, \vec{\lambda}_A(\vec{A})). \tag{4.93}$$

This implies, that the second part of Eqs. (4.91), i.e., $M'+1$ equations $\langle \hat{B}_w \rangle = B_w$ (which are just definitions of each of mean fields B_w) have identical solutions for both $\hat{H}_\lambda(\vec{A}, \vec{\lambda})$ (4.68) and the corresponding density operator $\hat{\rho}_\lambda(\vec{A}, \vec{\lambda})$ (4.29) as well as for $\hat{H}_{g\lambda}(\vec{A}, \vec{B}, \vec{\lambda}, \vec{l})$ (4.89) and the corresponding $\hat{\rho}_{g\lambda}(\vec{A}, \vec{B}, \vec{\lambda}, \vec{l})$. Therefore, those two MF Hamiltonians lead to the same optimal values of all mean-fields and Lagrange multipliers, hence also to identical equilibrium MF Hamiltonians, i.e.,

$$\hat{H}_{g\lambda}(\vec{A}_0, \vec{B}_0, \vec{\lambda}_0, \vec{l}_0) = \hat{H}_\lambda(\vec{A}_0, \vec{\lambda}_0). \tag{4.94}$$

However, in contrast to $\hat{H}_\lambda(\vec{A}, \vec{\lambda})$, $\hat{H}_{g\lambda}(\vec{A}, \vec{B}, \vec{\lambda}, \vec{l})$ is of the form (4.78), and we can apply to it the transformations (4.79). This yields MF Hamiltonian $\hat{H}^{(\sim)}_{g\lambda}(\vec{A}, \vec{B}, \vec{\lambda}, \vec{l})$ of the form (4.81), equivalent to (4.89). In order to use all the results obtained up to this point, it is enough to

re-label mean-fields, $B_w \to A_s$, $s = M+1, \ldots, M+M'+1$, $M+M'+1 \to M$. Nonetheless, for completeness, we provide explicit formulas in terms of B_w and l_w variables. We have

$$\hat{H}^{(\sim)}_{g\lambda}(\vec{A},\vec{B},\vec{\lambda},\vec{l}) = -\sum_{s=1}^{M}\lambda_s(\hat{A}_s - A_s) - \sum_{w=1}^{M'+1} l_w(\hat{B}_w - B_w) + W(\vec{A},\vec{B}) \tag{4.95}$$

with

$$W(\vec{A},\vec{B}) = B_{M'+1} + C_0(\vec{A}) + \sum_{s=1}^{M} C_s(\vec{A})A_s + \sum_{w=1}^{M'} G_w(\vec{A})B_w. \tag{4.96}$$

Now, we eliminate Lagrange multipliers λ_s and l_w using Eqs. (4.30), i.e., for each $(\vec{A},\vec{B}) \in \mathcal{D}_{\mathcal{AB}}$ we have[39]

$$\lambda_s = \lambda^{\star}_s(\vec{A},\vec{B}) = -\frac{\partial W_g(\vec{A},\vec{B})}{\partial A_s}, \qquad l_w = l^{\star}_w(\vec{A},\vec{B}) = -\frac{\partial W_g(\vec{A},\vec{B})}{\partial B_t}. \tag{4.97}$$

From Eqs. (4.97) we obtain $l^{\star}_w(\vec{A},\vec{B}) = -G_w(\vec{A})$, $l = 1,2,\ldots,M'$, and $l^{\star}_{M'+1}(\vec{A},\vec{B}) = -1$. Therefore we may define $\hat{H}^{(\star)}_{g\lambda}(\vec{A},\vec{B}) \equiv \hat{H}^{(\sim)}_{g\lambda}(\vec{A},\vec{B},\vec{\lambda}^{\star}(\vec{A},\vec{B}),\vec{l}^{\star}(\vec{A},\vec{B}))$ as

$$\hat{H}^{(\star)}_{g\lambda}(\vec{A},\vec{B}) = \sum_{s=1}^{M}\frac{\partial W_g(\vec{A},\vec{B})}{\partial A_s}(\hat{A}_s - A_s) + \sum_{w=1}^{M'+1} G_w(\vec{A})(\hat{B}_w - B_w) + W(\vec{A},\vec{B}), \tag{4.98}$$

The corresponding density operator is given by

$$\hat{\rho}^{(\star)}_{g\lambda}(\vec{A},\vec{B}) = \frac{\exp\big(-\beta(\hat{H}^{(\star)}_{g\lambda}(\vec{A},\vec{B}) - \mu\hat{N})\big)}{\mathrm{Tr}[\exp\big(-\beta(\hat{H}^{(\star)}_{g\lambda}(\vec{A},\vec{B}) - \mu\hat{N})\big)]}. \tag{4.99}$$

Finally, making use of $\hat{\rho}^{(\star)}_{g\lambda}(\vec{A},\vec{B})$ (4.99), we solve the remaining $M+M'+1$ self-consistency equations in order to determine the equilibrium values of all the $M+M'+1$ mean-fields, i.e., components of $\vec{A}_0$ and $\vec{B}_0$.

4.8.7 Present approach and formalism of References [162, 163, 164]

The formal results obtained so far allow us also to compare the present formalism with that of References [162, 163, 164]. A detailed discussion of the differences and similarities between those two approaches is outside the scope of this Thesis, and will be provided in a separate publication. However, for completeness, below we comment briefly on some formal aspects of both methods.

In Refs. [162, 163, 164], expectation values of certain operators, $\hat{O}_i$, i.e., $O_i = \mathrm{Tr}(\hat{\rho}_{app}\hat{O}_i)$ (Eq. (3.3) of [162]) with respect to the mixed states $\hat{\rho}_{app}$ (Eq. (3.1) of [162]) are considered. $\hat{\rho}_{app}$ serves as approximation to the intractable density operator $\hat{\rho}$ (Eqs. (2.8) of of [162]), constructed from $\hat{O}_i$ operators.

Note, that if we assume that the MF Hamiltonian, and consequently, the MF density operator we consider are of the single-particle character (i.e., bilinear in creation and/or annihilation operators), then, by applying Wick's theorem [136, 165] we are able compute (at least in principle) expectation value of each of the $\hat{O}_i$ operators. The case of single-particle MF Hamiltonians

[39] By $\mathcal{D}_{\mathcal{AB}}$ we denote the set of $M+M'+1$ dimensional vectors $(A_1, A_2, \ldots, A_M, B_1, \ldots, B_{M'+1})$, for which Hamiltonian (4.89) is well defined.

and density operators seems to be the most important from the point of view of applications of the MF theory in condensed matter physics, which is our main interest. However, single-particle character of the MF Hamiltonian is not necessary, and e.g. in [162], other types of operators are considered as well (cf. Section V of the latter Reference). Also within our approach, the single-particle character of the MF Hamiltonian has not been utilized as yet, and the results obtained so far remains valid also in the general case.

Assume now, that one of the $\hat{O}_i$ operators is some many-body Hamiltonian $\hat{H}_e$, and that the $\hat{\rho}_{app}$ (Eq. (3.1) of [162]) is constructed using single-particle operators only. To compare our approach with that of References [162, 163, 164], we may choose the initial MF Hamiltonian $\hat{H}(\vec{A})$ in a way that the following condition

$$\langle \hat{H}_e \rangle = \langle \hat{H}(\vec{A}) \rangle \equiv W(\vec{A}), \tag{4.100}$$

is fulfilled (cf. equation (4.80) of the present Thesis). In such a situation, density operator $\hat{\rho}_{app}$ (Eq. (3.7) of [162]) and the 'effective' MF Hamiltonian (Eq. (3.8) of [162]) seems to correspond to our $\hat{\rho}^{(\star)}_{g\lambda}$ (4.99) and $\hat{H}^{(\star)}_{g\lambda}$ (4.98), respectively ($\hat{H}^{(\star)}_{g\lambda}$ differs form formula (3.8) of [162] only by the terms proportional to unit operator, which cancel out in $\hat{\rho}^{(\star)}_{g\lambda}$). This observation suggests, that at least in certain limits, both approaches should lead to the same results.

Nonetheless, within our approach, $W(\vec{A})$ may be an arbitrary function of mean-fields, e.g. the expectation value of $\hat{H}_e$ with respect to a trial state, which is not a simple mean-field (uncorrelated) state. The latter case is crucial for the construction of MF approach to the t-J model, discussed in Part III of the present Thesis. It is unclear to us, if the same is achievable within the formalism of References [162, 163, 164].

4.8.8 Transformations (4.71) in presence of variational parameters of a non mean-field character

For completeness, now we are going to discuss, how the results obtained so far in this Section are modified in the presence of the $\vec{b}$ variables of a non mean-field character, introduced in Sub-subsection 4.4.1. In the present situation, the coefficients C_s, G_w, as well as the $\hat{H}_{e0}$ part of (4.67) may depend on $\vec{b}$ in a nontrivial way. Therefore, the optimal solution of (4.70), i.e., $\vec{\lambda} = \vec{\lambda}(\vec{A}, \vec{b})$ is also $\vec{b}$-dependent, and for $s = 1, \ldots, M$ we obtain

$$C^{\lambda}_s(\vec{A}, \vec{\lambda}(\vec{A}, \vec{b}), \vec{b}) = C_s(\vec{A}, \vec{b}) - \lambda_s(\vec{A}, \vec{b}) \equiv \mathcal{C}^{\lambda}_s(\vec{A}, \vec{b}). \tag{4.101}$$

Using the same arguments as previously, we see that for a given $\vec{A}$ and $\vec{b}$, transformations (4.71) applied to some MF Hamiltonian $\hat{H}_z(\vec{A}, \vec{b}) \equiv \hat{H}_{\lambda}(\vec{A}, \vec{\lambda}(\vec{A}, \vec{b}), \vec{b})$ lead to MF Hamiltonians, which may differ from each other only by the term proportional to unit operator. Therefore, all such MF Hamiltonians lead to the identical density operator $\hat{\rho}_z(\vec{A}, \vec{b})$. Again, the functional form of the latter is unambiguously defined by the choice of $\hat{H}_{e0}$, $\{\hat{A}_s\}_{s=1}^{M}$, $\{\hat{B}_w\}_{w=1}^{M'}$ operators, as well as of $G_w(\vec{A}, \vec{b})$ functions. Moreover, for transformations (4.77), in particular for (4.79), we obtain equivalent MF Hamiltonians, leading to identical $\mathcal{F}_z(\vec{A}, \vec{b}) \equiv \mathcal{F}(\vec{A}, \vec{\lambda}(\vec{A}, \vec{b}), \vec{b})$, hence also to identical $\vec{A}_0, \vec{\lambda}_0$ and $\vec{b}_0$, identical equilibrium MF Hamiltonian $\hat{H}_{\lambda 0}$ (4.37) and density operator $\hat{\rho}_{\lambda 0}$ (4.37).

Also, the arguments for equivalence of $\hat{K}_{\lambda}(\vec{A}, \vec{\lambda})$ (4.68) and $\hat{H}_{g\lambda}(\vec{A}, \vec{B}, \vec{\lambda}, \vec{l}) - \mu\hat{N}$ (4.89) remain unchanged in the presence of $\vec{b}$ variables. To see this, it is enough to notice that because $\vec{b}$ appears only in $\hat{H}(\vec{A})$ part of $\hat{K}_{\lambda}(\vec{A}, \vec{\lambda})$, Eqs. (4.28) have the same form for both $\hat{H}_{\lambda}(\vec{A}, \vec{\lambda}, \vec{b})$ (4.68) and $\hat{H}_{g\lambda}(\vec{A}, \vec{B}, \vec{\lambda}, \vec{l}, \vec{b})$ (4.89). However, for MF Hamiltonian of the form (4.78), the coefficients $\mathcal{C}^{\lambda}_s$ (4.101) for $s = 1, \ldots, M$ do not depend on $\vec{b}$, and only the term $W(\ldots)$ (4.80) may remain non-trivially $\vec{b}$-dependent.

4.8.9 Vector character of mean-fields and Lagrange multipliers

In the present Thesis, from the very beginning we have been using vector notation for both the mean-fields $\vec{A} = (A_1, A_2, \ldots, A_M)$ and the Lagrange multipliers, $\vec{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_M)$. Yet, so far we have not exploited the vector character of those quantities.

Frequently, it is convenient to use linear combinations of the original mean fields. Such transformation of the MF variables, denoted $\vec{X} = \vec{X}(\vec{A})$ (with the inverse transformation $\vec{A} = \vec{A}(\vec{X})$) is related to the corresponding transformation of $\hat{A}_s$ operators

$$\hat{X}_u = \sum_u \mathcal{U}^w_u \hat{A}_w \equiv \mathcal{U}^w_u \hat{A}_w, \qquad \hat{A}_s = \sum_t (\mathcal{U}^{-1})^t_s \hat{X}_t \equiv (\mathcal{U}^{-1})^t_s \hat{X}_t. \tag{4.102}$$

Obviously, (4.102) implies the same transformations for the mean-fields,

$$X_u = \sum_u \mathcal{U}^w_u A_w \equiv \mathcal{U}^w_u A_w, \qquad A_s = \sum_t (\mathcal{U}^{-1})^t_s X_t \equiv (\mathcal{U}^{-1})^t_s X_t. \tag{4.103}$$

Now let us consider the coefficients $C^\lambda_s(\vec{A}, \vec{\lambda})$ (4.69) appearing in $\hat{H}_\lambda(\vec{A}, \vec{\lambda})$ (4.68). Even without applying transformations (4.79), i.e., without assuming $\hat{H}_\lambda(\vec{A}, \vec{\lambda})$ of the form (4.78), we may formally solve Eqs. (4.30), which yields solution denoted here as $\vec{\lambda} \equiv \vec{\lambda}_v(\vec{A})$. It is easy to check that

$$C^\lambda_s(\vec{A}, \vec{\lambda}_v(\vec{A})) = \frac{\partial \langle \hat{H}(\vec{A}) \rangle}{\partial A_s} = \frac{\partial W(\vec{A})}{\partial A_s}. \tag{4.104}$$

In particular, this remains true for $\hat{H}^{(\star)}_\lambda(\vec{A})$ (4.83), for which we obtain $\vec{\lambda}_v(\vec{A}) = \vec{\lambda}^\star(\vec{A})$. After transformation (4.103), we have $W(\vec{A}) \to W_X(\vec{X}) \equiv W(\vec{A}(\vec{X}))$, and $\vec{\lambda} \to \vec{\lambda}_X$. The term $\vec{X} \cdot \vec{\lambda}_X$ transforms according to (we use Einstein summation convention)

$$\begin{aligned} X_u \frac{\partial W_X(\vec{X})}{\partial X_u} &= \mathcal{U}^w_u A_w \frac{\partial W(\vec{A}(\vec{X}))}{\partial A_s} \frac{\partial A_s}{\partial X_u} = \mathcal{U}^w_u A_w \frac{\partial W(\vec{A})}{\partial A_s} (\mathcal{U}^{-1})^u_s \\ &= \mathcal{U}^w_u (\mathcal{U}^{-1})^u_s A_w \frac{\partial W(\vec{A})}{\partial A_s} = \delta^w_s A_w \frac{\partial W(\vec{A})}{\partial A_s} = A_s \frac{\partial W(\vec{A})}{\partial A_s}. \end{aligned} \tag{4.105}$$

In above, δ^w_s is the Kronecker delta. We see, that in equilibrium, the mean-fields and the corresponding operators transform as contravariant vectors, whereas the Lagrange multipliers transform as covariant vectors.[40] Therefore, the equilibrium MF Hamiltonian is invariant with respect to transformations (4.103), as the scalar product-like notation of the terms like $\vec{X} \cdot \vec{\lambda}_X$ or $\vec{A} \cdot \vec{\lambda}$ suggests. The MF Hamiltonian is constructed only from such invariants (clearly, $W(\vec{A})$ also transforms as a scalar), cf. Eqs. (4.81) and (4.83).

4.9 Mean-field Hamiltonians of Hartree-Fock form

In the present Thesis we repeatedly refer to 'Hartree-Fock' (HF) type of MF Hamiltonians. Below we explain precisely what we mean by this term. In this Subsection, in order to construct MF Hamiltonian $\hat{H}(\vec{A})$, we use non-MF Hamiltonian $\hat{H}_e$ of the following form

$$\hat{H}_e = \hat{H}_{0e} + \sum_{\{\kappa,\gamma\}} V_{\kappa\gamma} \hat{\mathcal{O}}_\kappa \hat{\mathcal{O}}_\gamma. \tag{4.106}$$

We assume that each $\hat{\mathcal{O}}_\kappa$ operator is bilinear in creation and/or annihilation operators, and the summation is taken over all pairs $\{\kappa, \gamma\}$ of multi-indices.

[40] We do not take this 'tensor notation' too seriously, and write λ_s instead of λ^s.

We may decouple the interaction term in (4.106) (i.e., replace it by its MF counterpart) according to[41] (cf. Ref. [136])

$$\hat{\mathcal{O}}_\kappa \hat{\mathcal{O}}_\gamma \rightarrow \hat{A}_s A_t + A_s \hat{A}_t - A_s A_t. \tag{4.107}$$

$A_s = \langle \hat{A}_s \rangle$ and $A_t = \langle \hat{A}_s \rangle$ are given by (4.8) for the appropriate MF density operator, and $s = s(\kappa, \gamma)$, $t = t(\kappa, \gamma)$. In result, we obtain mean-field Hamiltonian of the following form

$$\hat{H}(\vec{A}) \;\; = \;\; \hat{H}^{HF}(\vec{A}) = \hat{H}_{0e} + \sum_{\{s,t\}} \tilde{V}_{st} \big(\hat{A}_s A_t + A_s \hat{A}_t - A_s A_t \big). \tag{4.108}$$

In above, $\hat{H}_{0e}$ does not depend on $\vec{A}$, and $\tilde{V}_{st}$ are linear combinations of matrix elements $V_{\kappa\gamma}$. Depending on the choice of $\hat{A}_s$ and $\hat{A}_t$ operators, $\hat{H}(\vec{A})$ of the form (4.108) is termed the *Hartree* or *Hartree-Fock* (HF) type MF Hamiltonian.

Within the framework of the present method, we add the constraint terms to $\hat{H}^{HF}(\vec{A})$, and obtain $\hat{H}_\lambda$ (4.15) of the form

$$\hat{H}^{HF}_\lambda(\vec{A}, \vec{\lambda}) = \hat{H}_{0e} + \sum_{\{s,t\}} \tilde{V}_{st} \big(\hat{A}_s A_t + A_s \hat{A}_t - A_s A_t \big) - \sum_s \lambda_s (\hat{A}_s - A_s). \tag{4.109}$$

Then, one can easily find that Eqs. (4.30) read now

$$\lambda_{t0} = -\Big\langle \frac{\partial \hat{H}^{HF}(\vec{A})}{\partial A_t} \Big\rangle_0 = -\Big\langle \sum_s \tilde{V}_{st} \big(\hat{A}_s - A_s \big) \Big\rangle_0, \quad t = 1, \ldots, M. \tag{4.110}$$

From Eqs. (4.31), with $\hat{H}(\vec{A})$ given by (4.108), and from (4.109) it follows that for any t, $\lambda_{t0} = 0$. Also, for Hamiltonian of such form, the solution of M self-consistent equations (4.8) or (4.31) coincide with the solutions of M equations $\nabla_A \mathcal{F} = 0$, $\vec{\lambda} = \vec{0}$ of the unconstrained variational approach. This is true, provided that the matrix $\tilde{V}_{st}$ is nonsingular (which can be always achieved by adding arbitrary small numbers ϵ_{st} to $\tilde{V}_{st}$), as then the only solution of the set of equations (4.110) is $\langle \hat{A}_s \rangle = A_s$, $s = 1, \ldots, M$ (a related discussion of this problem may be found in [136]). Therefore, for MF Hamiltonian of the Hartree-Fock type, application of our method for the equilibrium situation is not necessary. Still, it is required to construct the Landau potential $\mathcal{F}_z(\vec{A})$ and to analyze the non-equilibrium situation, cf. Section 4.5 and Supplement D (Subsection 12.1).

Obviously, MF Hamiltonian may be of the HF form with respect to only some of its mean-field variables. The equilibrium values of Lagrange multipliers corresponding to such variables are equal zero. The above discussion shows also, that the form (4.108) of $\hat{H}(\vec{A})$ seems to be a privileged one, because such choice automatically yields the self-consistent variational mean-field scheme without introducing any constraints. On the other hand, in the previous Subsection it has been shown, that within the framework of the present approach, apparently very different MF Hamiltonians are in fact completely equivalent, if only they are related by transformations (4.77). If we apply (4.77) to (4.108), for $\kappa \neq 1$ the resulting MF Hamiltonian will be no longer of the HF form. Consequently, unconstrained variational scheme will be no longer equivalent to that based on the BdG non-variational treatment, i.e., $\nabla_A \mathcal{F} = 0$, $\vec{\lambda} = \vec{0}$ is not equivalent to $\nabla_\lambda \mathcal{F} = 0$, $\vec{\lambda} = \vec{0}$. Different values of κ lead to different decoupling schemes,

[41] The $\hat{\mathcal{O}}_\kappa \hat{\mathcal{O}}_\gamma$ term may be decoupled in more than one way (which correspond to different possible contractions appearing in Wick's theorem), but this does not change the arguments presented here. Also, e.g. for the Bose-Hubbard model [166], it is not the interaction term which is decoupled according to (4.107), but the kinetic energy part. Still, the present conclusions remain valid for the latter case.

and none of them is privileged from the point of view of the present method. For example, instead of (4.107) one may postulate a decoupling scheme given by

$$\hat{\mathcal{O}}_\kappa \hat{\mathcal{O}}_\gamma \to \frac{1}{2}(\hat{A}_s A_t + A_s \hat{A}_t). \tag{4.111}$$

Note, that both (4.107) and (4.111) are symmetric in s and t, and both yield identical expectation value of the interaction term as a function of mean fields, i.e., $\langle \hat{\mathcal{O}}_\kappa \hat{\mathcal{O}}_\gamma \rangle = A_s A_t$. Additionally, both prescriptions avoid 'double counting' of the interactions. Therefore, if only the self-consistency of the variational approach is ensured by the presence of appropriate constraint terms, there is no obvious reason why (4.107) should be preferred over (4.111). As discussed in Subsection 4.8, within our approach the MF model with interaction term given by (4.111) yields results identical to those obtained for (4.107). Clearly, this would be not the case for the non-variational treatment based on BdG equations.

Summarizing, a construction of the MF Hamiltonian from its non- MF counterpart $\hat{H}_e$ is apparently not unique. However, the equivalence of different such obtained MF Hamiltonians is in fact guaranteed in the framework of our method. Clearly, this is a desired feature. It is reassuring, that the results do not depend on the choice of the decoupling scheme, which is quite arbitrary, but only on the set of operators used to construct MF Hamiltonian (which determines the universality class of this MF Hamiltonian, cf. Sec. 4.8), and on the form of the functional dependence of $W(\vec{A})$ (4.80) on $\vec{A}$ (which determines the equivalence class).

4.10 Relation of the present method to variational principle of Bogoliubov and Feynman

In the previous Subsections we have been considering a situation, when the description of the system is based solely on the MF Hamiltonian, in the sense, that no many-body non-MF Hamiltonian $\hat{H}_e$ has been required to determine the optimal values of mean-fields, although $\hat{H}_e$ may has been used to obtain the MF Hamiltonian. However, MF density operators are sometimes used as trial states within the variational principle based on the Bogoliubov-Feynman inequality [107, 108] and its generalizations [108]. Therefore, we would like to comment on the relation between the latter approach and the present MaxEnt-based method.

The Bogoliubov-Feynman inequality may be written as

$$\Omega_e \leq \langle \hat{K}_e - \hat{K}_a \rangle_a + \Omega_a = \mathrm{Tr}[\hat{\rho}_a(\hat{K}_e - \hat{K}_a)] + \Omega_a. \tag{4.112}$$

In the above, $\hat{H}_a$ is a trial ('approximate') Hamiltonian and $\langle \ldots \rangle_a \equiv \mathrm{Tr}[\hat{\rho}_a(\ldots)]$. We have also $\hat{K}_e = \hat{H}_e - \mu_e \hat{N}$, ($\hat{K}_a = \hat{H}_a - \mu_a \hat{N}$), with μ_e (μ_a) being the respective chemical potentials. Ω_e and Ω_a are the corresponding grand potentials,

$$\Omega_e = -\beta^{-1} \ln \mathrm{Tr}[e^{-\beta(\hat{H}_e - \mu_e \hat{N})}], \qquad \Omega_a = -\beta^{-1} \ln \mathrm{Tr}[e^{-\beta(\hat{H}_a - \mu_a \hat{N})}]. \tag{4.113}$$

To prove (4.112), one may use Klein inequality,

$$\mathrm{Tr}[\hat{\rho} \ln \hat{\rho}] \geq \mathrm{Tr}[\hat{\rho} \ln \hat{\sigma}], \tag{4.114}$$

which holds for arbitrary normalized density operators $\hat{\rho}$ and $\hat{\sigma}$ [186]. Inserting

$$\begin{aligned} \hat{\rho} &= \hat{\rho}_a = Z_a^{-1} e^{-\beta(\hat{H}_a - \mu_a \hat{N})}, \qquad Z_a = \mathrm{Tr}[\exp(-\beta(\hat{H}_a - \mu_a \hat{N}))], \\ \hat{\sigma} &= \hat{\rho}_e = Z_e^{-1} e^{-\beta(\hat{H}_e - \mu_e \hat{N})}, \qquad Z_e = \mathrm{Tr}[\exp(-\beta(\hat{H}_e - \mu \hat{N}))], \end{aligned}$$

into (4.114), we obtain (4.112). If one wishes to work with the chemical potential as an independent variable, then we have $\mu_e = \mu_a \equiv \mu$, and (4.112) reduces to

$$\Omega_e \leq \langle \hat{H}_e - \hat{H}_a \rangle_a + \Omega_a = \mathrm{Tr}[\hat{\rho}_a(\hat{H}_e - \hat{H}_a)] + \Omega_a. \tag{4.115}$$

However, usually it is more convenient to have particle number N as an independent variable. In that case the value of N is assumed to be the same both for the 'exact' and 'approximate' model. Also, μ_a and the unknown μ_e are not equal in general, as the former quantity is determined from different condition then the latter, i.e., $\langle \hat{N} \rangle_a = N$ and $\langle \hat{N} \rangle_e \equiv \mathrm{Tr}[\hat{\rho}_e \hat{N}] = N$, respectively. Inequality (4.112) may be then rewritten as

$$F_e = \Omega_e + \mu_e N \leq \langle \hat{H}_e - \hat{H}_a \rangle_a + \Omega_a + \mu_a N = \langle \hat{H}_e - \hat{H}_a \rangle_a + F_a. \tag{4.116}$$

Bogoliubov-Feynman inequality provide us with the upper bound for the exact grand potential Ω_e or free energy F_e of the system described by $\hat{H}_e$. The best trial state $\hat{\rho}_a$ is the one minimizing r.h.s. of (4.115), or (4.116), depending on the choice of independent thermodynamic variable, μ or N, respectively.

Let us consider now trial density operator $\hat{\rho}_a$ depending on some mean-field variables. For example, we may take $\hat{\rho}_a = \hat{\rho}(\vec{A})$ (4.11). However, if the mean-fields are treated as variational parameters in an unwary manner, there is no certainty that the basic self-consistency conditions (4.8) are satisfied. One may do not consider this fact as an important drawback, arguing that the only role of $\hat{\rho}_a$ is to provide the optimal (i.e., the lowest) upper bound for Ω_e or F_e. Therefore, one may not care too much about the internal consistency of the description based on such obtained MF density operator, and the latter is treated then as an auxiliary quantity. Still, in our opinion this point of view is unacceptable. Rather, the r.h.s. of (4.115) or (4.116) with $\hat{\rho}_a = \hat{\rho}$ given (4.11) may be minimized *only* with respect to the variables of a non-MF character, $b_1, \ldots, b_P$, if they are present in a problem at hand.[42] In such case, values of the mean-fields should be obtained in a self-consistent fashion through (4.8).

Still, mean fields can be treated as variational parameters, provided that we are able to ensure the self-consistency conditions (4.8). The simplest way to achieve this is to use the trial MF density operators of the form (4.43), with the trial Hamiltonian $\hat{H}_a = \hat{H}_z(\vec{A})$, yet bearing no relation to the $\hat{H}_e$ appearing in (4.115) or (4.116). Consequently, $\hat{\rho}_a = \hat{\rho}_z(\vec{A})$ (4.43) and then $\Omega_a = \Omega_a(\vec{A}) = \mathcal{F}_z(\vec{A}) = \mathcal{F}_z(T, V, \mu_a, \vec{h}; \vec{A})$ (4.44). For such choice of $\hat{\rho}_a$, conditions (4.8) are fulfilled automatically for any value of $\vec{A}$. Now, again we have to distinguish between the two cases, i.e., either we choose $\mu_e = \mu_a \equiv \mu$, or $\langle \hat{N} \rangle_a = \langle \hat{N} \rangle_e = N$; in the latter case we put $\mu_a \equiv \mu$. In the former case, we can rewrite (4.115) as

$$\Omega_e \leq \mathrm{Tr}\big[\hat{\rho}_z(\vec{A})\big(\hat{H}_e - \hat{H}_z(\vec{A})\big)\big] + \mathcal{F}_z(\vec{A}) \equiv \mathcal{B}_\mu(\beta, \mu, \ldots; \vec{A}), \tag{4.117}$$

whereas for $\langle \hat{N} \rangle_a \equiv \langle \hat{N} \rangle_z = \langle \hat{N} \rangle_e = N$, from (4.116) we obtain

$$F_e \leq \mathrm{Tr}\big[\hat{\rho}_z(\vec{A})\big(\hat{H}_e - \hat{H}_z(\vec{A})\big)\big] + \mathcal{F}_z(\vec{A}) + \mu A_1 \equiv \mathcal{B}_N(\beta, \mu, \ldots; \vec{A}) = \mathcal{B}_\mu(\ldots) + \mu A_1. \tag{4.118}$$

Note, that before the minimization procedure is carried through, both $A_1 = \langle \hat{N} \rangle_z$ (being usually one of the relevant mean-fields) and μ are independent variables of $\mathcal{F}_z$, $\mathcal{B}_\mu$ and $\mathcal{B}_N$. This remains in a full analogy to the situation encountered in Subsection 4.8.

To obtain the optimal upper bound for Ω_e or F_e, we minimize the r.h.s. of (4.117) or (4.118) with respect to mean fields. For $\mathcal{B}_\mu$, the $2M$ necessary[43] conditions for $\mathcal{B}_\mu(\beta, \mu, \ldots; \vec{A})$ to have a minimum acquire the form

$$\nabla_A \mathcal{B}_\mu(\beta, \mu, \ldots; \vec{A}) = 0. \tag{4.119}$$

[42] In what follows the presence of such variables is ignored for simplicity, but extension to the general case is straightforward.

[43] We assume, that the minimum corresponds to a stationary point.

On the other hand, minimization of $\mathcal{B}_N$ must be performed in a slightly different manner. Namely, due to the presence of $(-\mu A_1)$ term, the r.h.s. of (4.118) depends on μ and λ_1 only via $\tilde{\mu} = \mu + \lambda_1$. Therefore, we minimize $\mathcal{B}_N$ with respect to all mean-fields except A_1. We obtain

$$\frac{\partial \mathcal{B}_N(\beta, \mu, \ldots; \vec{A})}{\partial A_2} = 0, \qquad \ldots, \qquad \frac{\partial \mathcal{B}_N(\beta, \mu, \ldots; \vec{A})}{\partial A_M} = 0. \tag{4.120}$$

To find a relation between the solutions of Eqs. (4.119) and Eqs. (4.120), let us first note, that the self-consistency conditions (4.41) determine only $\tilde{\mu}(\vec{A})$, but not μ and $\lambda_1(\vec{A})$ separately. Therefore, the optimal values of the latter two variables cannot be found by solving (4.120). However, μ may determined from Eqs. (4.119); equation $\partial \mathcal{B}_\mu(\beta, \mu, \ldots; \vec{A})/\partial A_1 = 0$, not appearing in (4.120), can be solved either for μ or for N. Because μ is an $\vec{A}$-independent quantity (cf. related discussion in Section 4.7), the remaining equations of (4.119) and (4.120) are clearly identical. Therefore, if we use the same value of μ (e.g. determined from (4.119) for a given $A_1 = N$) in (4.119) and (4.120), the solutions of the former set of equations would be identical to those of the latter. The optimal solution, i.e. values of $\vec{A}$ and $\vec{\lambda}(\vec{A})$, for which the r.h.s. of either from (4.115) or (4.116) has the minimal value, will be denoted $\vec{A}_0^{BF}$ and $\vec{\lambda}_0^{BF} = \vec{\lambda}(\vec{A}_0^{BF})$. Remembering that $\hat{H}_z(\vec{A}) = \hat{H}_\lambda(\vec{A}, \vec{\lambda}(\vec{A}))$, we can rewrite (4.115) and (4.116) as

$$\Omega_e \leq \mathrm{Tr}\big[\hat{\rho}_\lambda(\vec{A}_0^{BF}, \vec{\lambda}_0^{BF})\big(\hat{H}_e - \hat{H}_\lambda(\vec{A}_0^{BF}, \vec{\lambda}_0^{BF})\big)\big] + \mathcal{F}(\vec{A}_0^{BF}, \vec{\lambda}_0^{BF}) \tag{4.121}$$

and

$$F_e \leq \mathrm{Tr}\big[\hat{\rho}_\lambda(\vec{A}_0^{BF}, \vec{\lambda}_0^{BF})\big(\hat{H}_e - \hat{H}_\lambda(\vec{A}_0^{BF}, \vec{\lambda}_0^{BF})\big)\big] + \mathcal{F}(\vec{A}_0^{BF}, \vec{\lambda}_0^{BF}) + \mu N. \tag{4.122}$$

The above discussion remain valid for other inequalities (i.e., various generalizations of (4.112)), and the corresponding variational principles built upon them, [108]. In such case, one may also use density operators $\hat{\rho}_z$ (4.43) as convenient variational *Ansätze*. Nonetheless, $\mathcal{S}_\lambda$ (4.14) together with the corresponding variational principle play a special role, as they determine the functional form (4.29) and (4.43) of the proper GC MF density operator. If the additional variational parameters of non-MF character, $b_1, \ldots, b_P$, are present, we simply minimize the r.h.s. of (4.115) with respect to each of b_l, this step contributes with the additional P equations appearing in Eqs. (4.119) or (4.120).

A hybrid approach leading to (4.119) and (4.120) is an attempt to find a compromise between two entirely different points of view on the MF approach. Therefore it cannot be regarded as fully satisfactory from either side. First, obviously, such approach yields higher upper bound for Ω_e or F_e, than the one resulting from the unconstrained variational procedure with no self-consistency conditions imposed. Also, note that the exact evaluation of $\langle \hat{H}_e - \hat{H}_a \rangle_a$ may be impossible in practice, but any approximate ways of evaluating this term may invalidate the Bogoliubov-Feynman inequality. Consequently, the whole formalism is less natural and transparent than the one founded on MaxEnt principle and based entirely on the MF Hamiltonian. In particular, natural definitions of the MF thermodynamic potentials as well as connection with Landau theory of phase transitions are obscured or even lost.

If $\mathcal{B}_\mu(\vec{A}) \neq \mathcal{F}_z(\vec{A})$, then usually also $\vec{A}_0^{BF} \neq \vec{A}_0$ and $\vec{\lambda}_0^{BF} \neq \vec{\lambda}_0$, with $\vec{A}_0$ and $\vec{\lambda}_0$ being the solutions of Eqs. (4.34). Consequently, $\hat{\rho}_z(\vec{A}_0^{BF}) = \hat{\rho}_\lambda(\vec{A}_0^{BF}, \vec{\lambda}_0^{BF})$ is no longer a true grand-canonical MF density operator for $\hat{H}_\lambda$. However, for any pair $(\hat{H}_e, \hat{H}_\lambda)$, for which

$$\langle \hat{H}_e - \hat{H}_a \rangle_a = \langle \hat{H}_e - \hat{H}_z \rangle_z = 0, \tag{4.123}$$

we have $\Omega_a(\vec{A}) = \mathcal{F}_z(\vec{A}) = \mathcal{B}_\mu(\vec{A})$. In such case the Bogoliubov - Feynman and MaxEnt variational principles coincide, and the MF density operator $\hat{\rho}_z(\vec{A}_0) = \hat{\rho}_\lambda(\vec{A}_0, \vec{\lambda}_0)$, which is optimal from the point of view of the MaxEnt inference, provides us also with the upper bound for exact grand potential Ω_e or free energy F_e of $\hat{H}_e$. This follows immediately from (4.119).

4.11 Formalism interpretations

Before we summarize this part of the Thesis, some remarks, both of a general, as well as of a more technical character, are in place here.

4.11.1 Time dependence of mean-field variables

At the early stage of the construction of the present formalism, we have assumed that both mean fields and Lagrange multipliers are time-independent quantities, and also that $[\hat{\rho}_\lambda, \hat{H}_\lambda] = 0$, cf. Eqs. (4.18). This condition allowed us to obtain an explicit form (4.29) of the density operator $\hat{\rho}_\lambda \equiv \hat{\rho}_\lambda(\vec{A}, \vec{\lambda}, \vec{b})$. Now we should make an *a posteriori* check, that the assumptions we have made, are consistent. Indeed, we may invoke Eq. (4.20), Eq. (4.8) with $\hat{\rho}_e \to \hat{\rho}_\lambda$, $\hat{H}_e \to \hat{K}_\lambda$, and $\hat{A} \to \hat{A}_s$ for each $s = 1, \ldots, M$, i.e.,

$$\begin{aligned} A_s(t) &= \sum_i \langle i(t)|\hat{\rho}_\lambda \hat{A}_s|i(t)\rangle = \sum_i \langle i(0)|e^{it\hat{K}_\lambda/\hbar}\hat{\rho}_\lambda \hat{A}_s e^{-it\hat{K}_\lambda/\hbar}|i(0)\rangle \\ &= \sum_i \langle i(0)|e^{-it\hat{K}_\lambda/\hbar}e^{it\hat{K}_\lambda/\hbar}\hat{\rho}_\lambda \hat{A}_s|i(0)\rangle = \sum_i \langle i(0)|\hat{\rho}_\lambda \hat{A}_s|i(0)\rangle \end{aligned} \tag{4.124}$$

In principle, depending on the MF model, there may exist solutions of Eqs. (4.124) for which $\partial\langle\hat{A}_s\rangle/\partial t = \partial A_s(t)/\partial t \neq 0$ and $\partial\lambda_s(t)/\partial t \neq 0$. If this was be the case, such solutions should be rejected within the present approach. However, it is always legitimate to assume constant $A_s(t) = A_s$ and $\lambda_s(t) = \lambda_s$, and then check if such assumption is consistent, namely if Eqs. (4.124) have any solution, i.e., if $\vec{A} \in \mathcal{D}_{\mathcal{A}}$. Then, both sides of (4.124) are obviously time - independent, provided that also $\partial\hat{A}_s/\partial t = 0$.

Therefore, we concentrate only on the stationary situation. This assumption is justified, if the time scales on which changes of order parameters occur are much larger then the characteristic time scale of the experiment. In such a situation we may assume that $\vec{A}$ appearing in $\hat{H}(\vec{A})$ is practically time-independent. Obviously, this ceases to be valid if the fluctuations are too strong. We also comment on this point in Supplement A (Subsection 12.1).

Apparently, it would be tempting to have at our disposal a MF approach which allows for a non-trivial time dependence of the non-equilibrium values of both $\vec{A}$ and $\vec{\lambda}$. For example, some mechanism, by which the system is driven to the thermodynamic equilibrium during the time evolution may be build into such formalism. However, compared to the present formulation, such MF theory would be much more complicated, and its construction is beyond the scope of this Thesis. Still, we may hope, that within a more complete time-dependent MF theory, description of the equilibrium situation will be very similar or even the same as the present one.

4.11.2 Super-selection rules

As mentioned in Section 3, an interpretation of mean-fields as semi-classical quantities imposes some restrictions on the superposition principle. Namely, it seems natural to postulate that the coherent superpositions of quantum states corresponding to different values of at least one of the mean-fields are forbidden, and that such states belong to different and mutually orthogonal Hilbert spaces. In other words, we postulate the existence of certain super-selection rules. Similar situation is encountered in some non-MF models, but only in the thermodynamic limit. For example, in the absence of the external field, the (degenerate) ground state of the Heisenberg ferromagnet [147] (as well as all of its states having finite number of excitations, i.e., finite number of reversed spins), related to different directions of the magnetization, are orthogonal. In contrast, for the MF models, numerical value of a scalar product of two quantum

states, each of which is characterized by different values of mean-fields, does not have to be (and is usually not) equal to zero. However, if additionally the model is supplemented with the above provided *interpretation*, then for a finite system we may obtain a situation similar to that encountered in the thermodynamic limit for non-MF Hamiltonians.

4.11.3 Internal consistency of mean-field approach

MF Hamiltonian $\hat{H}(\vec{A})$ (4.1), the self-consistent Hamiltonian $\hat{H}_z(\vec{A})$ (4.42) and the related MF density operator $\hat{\rho}_z(\vec{A})$ (4.43) are operator-valued function of the mean-fields $A_1, \ldots, A_M$, i.e., they are characterized by a *single*, well defined $\vec{A} = (A_1, A_2, \ldots, A_M)$ - there is no averaging over different values of $\vec{A}$. This simple fact has important consequences for the properties of the present MF formalism and its physical interpretation.

As pointed out above, when treating mean-fields as semi-classical objects, we should not make coherent superpositions of the states differing by $\vec{A}$ or $\vec{\lambda}$. In fact, we have to exclude any, either coherent or statistical, mixing of the states characterized by a macroscopically distinguishable values of $\vec{A}$ (of course, we have to define precisely what we mean by 'macroscopically distinguishable'). Therefore, to assure internal consistency of the MF description, each A_s should correspond to a weakly-fluctuating physical quantity. Only if the latter condition is fulfilled, application of the MF approach in the present form is legitimate. This, in turn, imposes certain restrictions on the choice of $\hat{A}_s$ operators. This problem will be discussed also in Supplement A (Subsection 12.1).

4.11.4 Physical and statistical aspects of MF statistical mechanics

According to Jaynes, the only place where the non-MF statistical mechanics makes contact with the laws of physics, is the choice of a microscopic Hamiltonian. The remaining part of the standard statistical-mechanical formalism is just the maximum-entropy inference, 'which principles are independent of any physical properties' [109]. A distinction between 'physical' and 'statistical' aspects of the MaxEnt-based statistical mechanics can also be made in the case of the mean-field approach. However, such distinction is not as sharp as in the non-MF case. Here, the 'physical' part contains also the choice of the mean fields $\vec{A}$, which are relevant for the problem at hand, and therefore required in order to construct the MF Hamiltonian. Nonetheless, equilibrium values of mean fields ($\vec{A} = \vec{A}_0$) and Lagrange multipliers ($\vec{\lambda} = \vec{\lambda}_0$), and consequently, the equilibrium MF grand Hamiltonian, $\hat{K}_{\lambda 0} = \hat{K}_\lambda(\vec{A}_0, \vec{\lambda}_0, \vec{b}_0)$ (4.36) are determined using MaxEnt principle, so they depend on the 'statistical' aspects of the formalism. Consequently, eigenstates of $\hat{K}_{\lambda 0}$ (being pure quantum states) may also depend on temperature, chemical potential and other quantities of purely statistical origin. This is legitimate, as long as the formulation is explicitly time-independent.

Note also, that within the standard statistical mechanics, some interaction with the environment is usually necessary in order to ensure thermalization. It is also assumed, that such interaction is weak enough and does not modify energy levels of the system, and therefore it is not explicitly present in the Hamiltonian [109, 112]. Also in the case of a finite-temperature mean-field description of the system, some interaction with the environment must be assumed for the same reason as in the standard situation. Yet, in contrast to the former case, here also the quantum micro states of the system are affected by such interaction.

It has been claimed that for $\vec{A} = \vec{A}_0$, $\vec{\lambda} = \vec{\lambda}_0$ and $\vec{b} = \vec{b}_0$, a density operator $\hat{\rho}_{\lambda 0}(\beta, \mu)$ (4.37), i.e., $\hat{\rho}_\lambda(\beta, \mu; \vec{A}_0, \vec{\lambda}_0, \vec{b}_0)$, is a proper grand-canonical (GC) MF density operator corresponding to MF Hamiltonian $\hat{H}(\vec{A}, \vec{b})$. This is not the case for $\hat{\rho}_{sc}^{(0)}$ (4.40), which has only an apparent GC form. Here we would like to provide further comments on this issue.

If we choose a concrete value of $\langle\hat{H}_\lambda\rangle = \langle\hat{H}\rangle = E$ and $\langle\hat{N}\rangle = N$, we have to determine Lagrange multipliers β and μ appearing in $\hat{\rho}_{\lambda 0}(\beta,\mu)$ (say $\beta = \beta_\lambda$, $\mu = \mu_\lambda$) and $\hat{\rho}^{(0)}_{sc}(\beta,\mu)$ ($\beta = \beta_{sc}$, $\mu = \mu_{sc}$), respectively. Obviously, in general case, $\beta_\lambda \neq \beta_{sc}$ and $\mu_\lambda \neq \mu_{sc}$. From the way both probability distributions are constructed, it should be clear, that for S_{vN} (4.4) we have

$$S_{vN}\big(\hat{\rho}_{\lambda 0}(\beta_\lambda,\mu_\lambda)\big) \geq S_{vN}\big(\hat{\rho}^{(0)}_{sc}(\beta_{sc},\mu_{sc})\big), \tag{4.125}$$

with equality for the density matrices of the Hartree-Fock MF models. Also, apart from $\hat{\rho}^{(0)}_{sc}(\beta_{sc},\mu_{sc})$, any other density operator leading to the same $\langle\hat{H}\rangle$, obeys inequality (4.125).

The point is that true GC MF density operator, $\hat{\rho}_{\lambda 0}(\beta_\lambda,\mu_\lambda)$ (4.37) (from now on denoted as $\hat{\rho}_{\lambda 0}(\beta,\mu)$) establishes one-to-one correspondence between E and β, similarly as in the case of a standard (non-MF) statistical mechanics [109]. Therefore, we can make the Legendre transformation $E \leftrightarrow \beta$, and work with a fixed β and variable E instead of fixed E and variable β. However, In such a situation, depending on the choice of N or μ as an independent variable, we select the optimal density operator by comparing values of free energy F or grand potential $\Omega = \mathcal{F}(\vec{A}_0,\vec{\lambda}_0,\vec{b}_0)$ (4.47), and not of the entropy. Clearly, for fixed β, $\hat{\rho}_{\lambda 0}(\beta,\mu)$ and $\hat{\rho}^{(0)}_{sc}(\beta,\mu)$ may yield different values of E.

The above conclusions are valid both when $[\hat{H}(\vec{A}),\hat{H}_\lambda(\vec{A},\vec{\lambda})] = 0$, i.e., when $\hat{H}$ and $\hat{H}_\lambda$ have common eigenbasis, as well as in the opposite case, when application of MaxEnt principle modifies algebraic structure of the model, and when the common eigenbasis of $\hat{H}$ and $\hat{H}_\lambda$ no longer exists.

4.11.5 Generalized entropies

Originally, MaxEnt principle has been founded on Shannon entropy (in classical case) or on the von Neumann entropy (in quantum case) [109]. However, MaxEnt principle is sometimes formulated in a more general form, which makes use of other entropies [187]. In particular, *Tsallis entropy* [188], leading to the so-called non-extensive statistical mechanics, gained a considerable interest in the last decades. The mean-field statistical mechanics may be also based on one of those alternative definitions of entropy, and this step requires only a small modification of the present formalism.

4.11.6 Lagrange multipliers as molecular fields

The form of the constraint terms (4.13) suggests an interpretation of Lagrange multipliers as a kind of molecular fields, coupled to the corresponding mean-fields. However, as shown in Section 4.8, apparently very different MF Hamiltonians are in fact completely equivalent, and lead to the same values of all physical quantities that can be computed within a given MF model.

In contrast to mean-fields, Lagrange multipliers acquire *different* equilibrium values for different equivalent MF Hamiltonians. Only certain combinations of λ_s with the mean-fields A_s are invariant, i.e., $C^\lambda_s(\vec{A},\vec{\lambda})$ (4.69) have the same equilibrium value for all equivalent MF Hamiltonians. This bears some analogy with the gauge-invariance; only the gauge-invariant quantities may be given physical interpretation. Similarly, here the only invariant quantities correspond to observables, e.g. the quasi-particle energies. Equilibrium values of Lagrange multipliers depend not only on the physical contents of the model but also on its particular formulation (out of many possible equivalent formulations).

5 Summary and discussion of Part II

To summarize this part of the Thesis, we have presented a method of solving mean-field (MF) lattice models. Our approach, founded on the maximum-entropy (MaxEnt) principle, has been devised for the situation, when description of the system is based entirely on the MF Hamiltonian. For that purpose, the original formulation of MaxEnt principle must be suitably generalized. Namely, both the probabilities (diagonal elements of the density operator), as well as the mean-fields (and possibly also other variables of a non-MF character) are treated as variational parameters. This step requires an introduction of the additional constraints, which, first, ensure the self-consistency of the mean-field model, and second, allow to treat mean-field averages and probabilities as independent variables. By *self-consistency* we understand that the optimal values of mean-fields, obtained from the variational procedure, are also expectation values of the corresponding operators, evaluated by using MF density operator. For an arbitrary MF model, this is not guaranteed *a priori*. Quite contrary, variational procedure which consists of minimization of an appropriate MF thermodynamic potential leads to self-consistent results only for important, but narrow class of Hartree-Fock MF Hamiltonians.

Modification of the standard MF formalism, which has been proposed here, is fairly straightforward, but carries out important consequences. First, from the point of view of Bayesian mathematical statistics, our method represents an optimal way of solving MF models. Second, a natural connection between microscopic MF models and Landau theory of phase transitions is established. On the other hand, within the standard formulation of the mean-field approach, the requirements of self-consistency are not fulfilled for non-equilibrium values of mean-fields. This, strictly speaking, does not allow for a consistent interpretation of the results of MF models in terms of Landau theory, even if this fact is sometimes ignored.

It also becomes possible to analyze the internal limitations of the present MF formalism, and to estimate, at least in the qualitative manner, the range of its applicability (this is discussed in Supplement A, i.e., Subsection 12.1).

The underlying variational principle guarantees that the thermodynamic quantities given by the first derivatives of the equilibrium free energy (or grand potential), i.e., the entropy, the pressure or the average particle number, are identical with their statistical-mechanical definitions based on the MF density operator. However, this can no longer hold for the second derivatives of thermodynamic potentials, i.e., specific heat, magnetic suspectibility, etc., where additional terms of the mean-field origin appear. Again, similar situation is encountered both in the Landau theory and microscopic mean-field theories (e.g. BCS theory).

Finally, let us note, that our approach provides also a simple and unique way of constructing the optimal MF approximation for a given non-MF (many-body) Hamiltonian. For a wide class of many-body Hamiltonians, this clarification may be achieved by invoking, apart from the present formalism, only Wick's theorem.

Part III
Mean-field theory of t-J model

6 t-J model

In this Section we discuss the t-J model, describing correlated fermions in a single narrow band, and analyze its different forms. Its relationship to the Hubbard model, the relevance for the physics of the high-temperature superconducting cuprates as well as its possible extensions and generalizations, are also discussed. This Chapter represents the principal application of the original formal method presented in Part II to the model composing the title of the present Thesis.

6.1 t-J Hamiltonian

Hamiltonian of the t-J model [8, 22, 23, 24, 25, 26, 27, 28, 30, 31, 33] reads

$$\hat{H}_{tJ} = \hat{P}\Big\{ \sum_{ij\sigma} t_{ij} c^{\dagger}_{i\sigma} c_{j\sigma} + \sum_{\langle ij \rangle} J_{ij} \left(\hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j - \frac{c_1}{4} \hat{n}_i \hat{n}_j \right) + c_2 \hat{H}_3 \Big\} \hat{P}. \tag{6.126}$$

The first term, $\hat{H}_t \equiv \hat{P} \sum_{ij\sigma} t_{ij} c^{\dagger}_{i\sigma} c_{j\sigma} \hat{P}$, is the projected kinetic (band) energy part, the second

$$\hat{H}_J \equiv \sum_{\langle ij \rangle} J_{ij} \Big(\hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j - \frac{c_1}{4} \hat{n}_i \hat{n}_j \Big), \tag{6.127}$$

expresses the kinetic exchange interaction, whereas the third ($\hat{H}_3$) is the so-called *correlated hopping* or *three-site* term. By $\sum_{ij}$ we denote the double summation over sites, with $i \neq j$, whereas by $\sum_{\langle ij \rangle}$ summation over the nearest-neighbor bonds (each bond is counted once). Coefficients c_1 and c_2, equal either 0 or 1, allow us to select a particular limiting form of the t-J model. The Gutzwiller projection operator (projector)

$$\hat{P} = \prod_i (1 - \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}), \tag{6.128}$$

eliminates double occupied sites. An explicit form of $\hat{H}_3$ reads

$$\hat{H}_3 = \sum_{ijk} \sum_{\sigma} \frac{t_{ij} t_{jk}}{U} \left(b^{\dagger}_{i\sigma} \hat{S}^{\bar{\sigma}}_j b_{k\bar{\sigma}} - b^{\dagger}_{i\sigma} \hat{n}_{j\bar{\sigma}} b_{k\sigma} \right), \tag{6.129}$$

with $b_{i\sigma} \equiv (1 - \hat{n}_{i\bar{\sigma}}) c_{i\sigma}$, $\hat{\nu}_{i\sigma} \equiv (1 - \hat{n}_{i\bar{\sigma}}) \hat{n}_{i\sigma}$, $\hat{\nu}_i = \sum_{\sigma} \hat{\nu}_{i\sigma}$, and $\hat{S}^{\sigma}_i \equiv b^{\dagger}_{i\sigma} b_{i\bar{\sigma}}$. Hamiltonian (6.126) in its complete form, i.e., with $c_1 = c_2 = 1$ can be recast to the equivalent, more compact and intuitive form involving real space pairing operators $\hat{B}_{ij}$ (cf. Ref. [27]), defined as

$$\hat{B}_{ij} \equiv \frac{1}{2} \left(b_{i\uparrow} b_{j\downarrow} - b_{i\downarrow} b_{j\uparrow} \right). \tag{6.130}$$

Namely, we can combine both two- and three- site interactions to obtain

$$\hat{H}_{tJ} = \Big(\sum_{ij\sigma} t_{ij} b^{\dagger}_{i\sigma} b_{j\sigma} - \sum_{ijk\sigma} \frac{2 t_{ij} t_{jk}}{U} \hat{B}^{\dagger}_{ij} \hat{B}_{jk} \Big). \tag{6.131}$$

Note, that due to the presence of $\hat{P}$ in (6.126), $\hat{H}_3$ (6.129) and $\hat{B}_{ij}$ operators may be equivalently expressed in terms of bare electron operators $c_{i\sigma}$ instead of $b_{i\sigma}$, etc. $\hat{H}_{tJ}$ (6.126) with $c_1 = c_2 = 1$ has been derived [22] from the single-band Hubbard Hamiltonian [19, 20, 21, 190]

$$\hat{H}_{tU} = \sum_{ij\sigma} t_{ij} c^{\dagger}_{i\sigma} c_{j\sigma} + U \sum_{i} \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}, \tag{6.132}$$

where $U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}$ is the on-site Coulomb repulsion term. Namely, by applying canonical transformation (CT) [22, 23, 24, 25, 191, 192] to $\hat{H}_{tU}$, i.e.,

$$\hat{H}_{tU} \rightarrow e^{-i\hat{S}} \hat{H}_{tU} e^{i\hat{S}} \equiv \hat{H}_{CT}, \tag{6.133}$$

with $\hat{S}^{\dagger} = \hat{S}$, expanding $e^{\pm i\hat{S}}$ and retaining the terms up to k-th order (up to $\hat{S}^k$), we obtain approximate form of $\hat{H}_{CT}$,

$$\hat{H}_{CT} \rightarrow \hat{H}^{(k)}_{tJ} \equiv \Big(1 - i\hat{S} + \ldots + \frac{(-i)^k}{k!} \hat{S}^k\Big) \hat{H}_{tU} \Big(1 + i\hat{S} + \ldots + \frac{(i)^k}{k!} \hat{S}^k\Big). \tag{6.134}$$

$\hat{H}^{(k)}_{tJ}$ is an extended t-J model, in which only the terms up to those of order $|t|^k/U^{k-1}$ appear.[44] The resulting higher-order terms are either the so-called *ring-exchange terms* [190, 193] or have a form similar to that of $\hat{H}_3$, though more complicated [25]. Higher-order corrections change also values of J_{ij} parameters, as they include additional contributions to the two-site virtual hopping.

In the $|t_{ij}| \ll U$ limit, one may keep in $\hat{H}_{CT}$ only the terms proportional to t_{ij}, $|t_{ij}|^2/U$ and $t_{ij}t_{jk}/U$, and obtain $\hat{H}^{(2)}_{tJ}$, i.e., the full t-J Hamiltonian (6.126) with $c_1 = c_2 = 1$ and $J_{ij} = 4|t_{ij}|^2/U$ [22, 23, 27, 28]. On the other hand, Hamiltonian (6.126) with $c_1 = c_2 = 0$, i.e., of the following simplified form

$$\hat{H}_{tJ} \equiv \hat{H}_t + \hat{H}_J = \hat{P} \left\{ \sum_{i,j,\sigma} t_{ij} c^{\dagger}_{i\sigma} c_{j\sigma} + \sum_{\langle ij \rangle} J_{ij}\ \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j \right\} \hat{P}. \tag{6.135}$$

has been proposed in Ref. [29] as an effective model describing the low-energy physics of the cuprates. Both (6.135), or (6.126) with $c_1 = 1$, $c_2 = 0$ are used to study high-T_c superconductivity, and three-site terms are usually disregarded. However, the role of $\hat{H}_3$ has not been fully elucidated [8], and some Authors (cf. e.g. [61, 194, 195]) include also this term. This issue is discussed in detail below.

Summarizing, the t-J model may be regarded as an effective Hamiltonian, describing the properties of the Hubbard model with large U at low energies. Due to the exclusion of doubly occupied sites, the Hilbert space dimension D_H is equal 3^{Λ} (strongly correlated metal), and $D_H = 2^{\Lambda}$ for the limiting case of the Mott-Hubbard insulator (with Λ being a number of lattice sites) as compared to $D_H = 4^{\Lambda}$ for the original Hubbard model.

6.2 General remarks on t-J model

6.2.1 Effective Hamiltonians: a broader perspective

According to Ref. [191], canonical transformation (6.133) may be regarded in an analogy to the Foldy - Wouthuysen (FW) transformation [196], used to derive the effective Schrödinger-Pauli equation from the Dirac equation. The effective Hamiltonian obtained in this manner

[44] With $t = t_{ij}$.

from Dirac Hamiltonian, apart from the local external vector and scalar potentials, contains also terms which are nonlocal, i.e., which involve derivatives of the scalar potential, e.g. the Zeeman, the spin-orbit and the Darwin terms. In the present case, the non-local (namely, involving at least two sites) terms, like kinetic exchange or $\hat{H}_3$, emerge from purely local (on-site) Coulomb interaction. This is the price to pay for dealing with the Hilbert space of reduced dimensionality. On the other hand, in both cases a transparent physical interpretation may be ascribed to the effective Hamiltonian resulting from CT. For example, for half-filled lattice ($n = 1$), in the $t \ll U$ limit, charge fluctuations are frozen, and the Hamiltonian (6.126) reduces to the Heisenberg Hamiltonian ($\hat{H}_{tJ} \equiv \hat{H}_J$ (6.127) with $c_1 = 1$ and $c_2 = 0$), which is a minimal model of correlated magnetic insulators [8, 190, 197, 198]. From the form of $\hat{H}_J$ it should be clear, that in the limit in question, the original electron system behaves like a collection of interacting spins (charge degrees of freedom are frozen out). This can be deduced in much more difficult manner by a direct analysis of the original Hubbard Hamiltonian (6.132).

6.2.2 t-J model as a minimal model of cuprate superconductors

As mentioned in the Introduction, $\hat{H}_{tJ}$, even in its incomplete form, i.e., either (6.135) or (6.126), with $c_1 = 1$ and $c_2 = 0$, is frequently regarded as a minimal electronic model of the high T_c compounds [8, 30, 31, 33, 39, 61]. Namely, this Hamiltonian contains all the necessary ingredients required to explain the physics of a lightly doped Mott-Hubbard insulator. First, the correlated hopping part $\hat{H}_t$ is responsible for a charge (hole) transport for a nonzero doping. Next, the kinetic exchange term (6.127) is regarded as the cause of the antiferromagnetism in the undoped parent compound and at low doping level; here also as a source of the *real-space* pairing. Indeed, within the mean-field approach both antiferromagnetism and superconductivity emerge in a natural manner from the presence of this particular term. Finally, the three-site term $\hat{H}_3$ may be interpreted as a kinetic energy operator for real-space pairs of electrons in a spin-singlet state, described by $\hat{B}_{ij}$ operators [27].

6.2.3 Nontrivial role of higher-order terms

Note, that even if the reasonable value of $|t_{ij}|/U$ in the cuprate compounds is 1/10 or smaller, it is not obvious that we can neglect the higher-order terms appearing in $\hat{H}_{tJ}^{(k)}$. Some of those terms disappear for half-filling ($n = 1$), and their influence is weak for small doping. Indeed, within the MF approach, many terms contain $(1-n)^{k'}$ pre-factors, where $k' \leq k$. However, even if such terms are small for $n \approx 1$, they can still affect behavior of the system at larger doping, e.g. near the upper critical concentration $n_c = 1 - x_c \approx 0.7 - 0.8$. Also, the number of higher-order terms in $\hat{H}_{tJ}^{(k)}$ is usually large [24, 25]. Even if the terms of a similar structure frequently appear with the opposite signs, and therefore their influence is weakened, a net effect can still be non-negligible.

Moreover, after the MF decomposition of $\hat{H}_{tJ}^{(k)}$ is made, many different terms in the resulting MF Hamiltonian are usually related to each term of $\hat{H}_{tJ}^{(k)}$. Such terms may partly cancel each other, but it is hard to estimate the overall effect without detailed calculations. In other words, factors of combinatorial origin may overcome the smallness of the $|t_{ij}|/U$ parameter as well as the influence of the $(1-n)^{k'}$ pre-factors.

According to Ref. [193], the higher-order terms (in particular, the ring-exchange terms), are important for a successful description of the undoped and lightly doped high-T_c compounds. If such point of view was accepted, then all the terms up to at least $\sim |t|^4/U^3$ should be included in an extended t-J Hamiltonian ($\hat{H}_{tJ}^{(4)}$). Next, one may perform the MF analysis of $\hat{H}_{tJ}^{(4)}$, at least for the simplest MF states. It would be interesting to check if the respective MF solutions have a proper asymptotic behavior with increasing order k of the canonical transformation (6.134).

This would make the whole MF-approach for the t-J model more formally correct. On the other hand, there is no certainty, that such procedure 'converges', i.e., that the smallness of $|t|/U$ factor eventually overcomes the influence of the large number of terms in the MF Hamiltonian. Obviously, the inclusion of 'fourth-order' ($\sim |t|^4/U^3$) and omission of the higher-order terms may still be insufficient.

As pointed above, the MF analysis of the higher-order terms is very tedious and to our knowledge, has never been fully carried through. Also in this Thesis, our aim is to compare the results of MF treatment of both $\hat{H}_{tJ}$ (6.126) with $c_1 = 1$, $c_2 = 1$, as well as its simplified versions: $c_1 = 0$, $c_2 = 0$ and $c_1 = 1$, $c_2 = 0$. Therefore, in what follows we ignore $\hat{H}_{tJ}^{(k)}$ for $k > 2$ altogether.

6.2.4 Mean-field treatment of Hubbard model

The above discussion shows that it is not completely clear, which particular form of the t-J model or its higher-order extensions should be used. Also, MF treatment of a more complicated versions the t-J model may be very tedious. Moreover, there are certain issues, which are obviously beyond the domain of applicability of the t-J model (i.e., the interaction-tuned Mott-Hubbard transition).[45] Also, on the MF level dimensionality of the Hilbert space (D_H) is not a limiting factor, and with this respect the t-J model has no advantages over the Hubbard model. Therefore, one may ask why we do not simply use the latter model instead of the former? The answer to this question is direct. Namely, in contrast to the t-J model, no simple MF-treatment of the Hubbard Hamiltonian with positive $U0$ leads in a natural manner to the pairing of required d-wave symmetry and a stable superconducting solution. On the technical level, this is caused by the absence of the exchange term in the starting Hubbard model. Therefore a MF description of the superconducting state in the cuprates based on the t-J model is preferred.

6.2.5 t-J-U model

Apart from the Hubbard and the t-J models, the hybrid t-J-U model, defined by the following Hamiltonian

$$\hat{H}_{tJU} = \sum_{ij\sigma} t_{ij} c_{i\sigma}^{\dagger} c_{j\sigma} + \sum_{\langle ij \rangle} J_{ij}\ \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}, \tag{6.136}$$

has been proposed by Zhang [200], cf. also [201]. In contrast to t-J Hamiltonian, $\hat{H}_{tJU}$ cannot be rigorously derived from any standard microscopic model of correlated fermions. However, it has been argued (cf. e.g. [202]), that on the MF level, phenomenologically motivated t-J-U model provides a better description of correlated systems than either the Hubbard or the t-J model. Note, that the MF treatment of $\hat{H}_{tJU}$ can be formulated in a similar manner to that of $\hat{H}_{tJ}$.

6.2.6 Other possible extensions of t-J model

The form of the t-J model, regarded as the derivative of the Hubbard model, depends crucially on the details of the Hubbard Hamiltonian. The point is, that single-band Hubbard model (6.132) may be by itself generalized in several ways. Namely, inter-site interactions of various kind, e.g. density-density Coulomb interaction, direct exchange, and more exotic correlated (*Coulomb-assisted*) hopping and pair-hopping terms may be also included [190, 203, 204]. In

[45] Undoped ($n = 1$) cuprates are well in the Mott-insulating regime, and we do not have to worry about the issue the interaction-tuned Mott-Hubbard transition. However, this may be important for other systems, e.g. organic superconductors [4].

such a situation, an effective model derived from such extended Hubbard model in the limit of strong correlations may significantly differ from (6.126). It should be noted, that the additional terms in the Hubbard Hamiltonian, for example inter-site Coulomb interaction, may be of comparable magnitude to those which are present in Hamiltonian (6.126). Therefore, such terms may be important for the physics of strongly correlated systems, cuprate superconductors in particular. Analysis of the such extended t-J model is also beyond the scope of the present Thesis (for a simple estimate of the role of intersite Coulomb interaction see [204].

7 Renormalized mean-field theory (RMFT)

At the beginning of this Section, we discuss quantum states, which are candidates for the approximate ground state of the t-J model. We show, that particular class of such variational trial states leads to an effective mean-field description, called the *renormalized mean-field theory* (RMFT).

First, we analyze RMFT in its standard formulation, i.e., in the case when the so-called Gutzwiller renormalization factors g^t and g^J can be defined. Next, this formulation is applied only to the simplest form of the t-J Hamiltonian (6.135), i.e., (6.126) with $c_1 = 0$ and $c_2 = 0$.

Our second aim here is to compare some of the various renormalization schemes proposed in the literature and eventually, to select the optimal one. In Section 7.4.3 we provide an alternative formulation of the RMFT approach, based on our original approach of Section 4.8.

7.1 Concept of resonating valence bond (RVB) state and correlated variational wave functions

In the preceding Section, the t-J model has been introduced as a purely electronic microscopic model of high-T_c cuprate superconductors. Unfortunately, exact solutions of this model are limited to very special cases. This is both due to the interaction terms, as well as to the presence of the projection operator $\hat{P}$. Even in the $J = 0$ limit, $\hat{H}_{tJ}$ (6.135) is not an independent-particle Hamiltonian. In such a situation, one may try to search for a proper trial wave function, being an acceptable approximation to the ground state of the t-J model. Such variational state has been proposed by Anderson [38, 39] in the form

$$|\Psi\rangle \equiv \hat{P}|\Psi_0\rangle = \prod_i (1 - \hat{n}_{i\uparrow}\hat{n}_{i\downarrow})|\mathrm{BCS}\rangle. \tag{7.137}$$

In above, $|\mathrm{BCS}\rangle$ is the Bardeen-Cooper-Schrieffer (BCS)-type state [41], i.e.,

$$|\mathrm{BCS}\rangle = \prod_{\mathbf{k}} (u_{\mathbf{k}} + v_{\mathbf{k}} c^{\dagger}_{\mathbf{k}\uparrow} c^{\dagger}_{-\mathbf{k}\downarrow})|0\rangle, \tag{7.138}$$

with $|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 = 1$. Explicit form of $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$, being the Bogoliubov coherence factors in the paired state, depends on the manner in which the BCS state is constructed. $|\Psi\rangle$ of the form (7.137) is termed *resonating valence bond* (RVB) state and denoted $|\mathrm{RVB}\rangle$. To keep our discussion more general, we will consider trial states of the following form

$$|\Psi\rangle = \hat{P}_C|\Psi_0\rangle. \tag{7.139}$$

Uncorrelated state $|\Psi_0\rangle$ is an eigenstate of a noninteracting, single-particle Hamiltonian $\hat{H}_R$ (not yet specified). $|\Psi_0\rangle$ may be of a more complicated form then the simple BCS state. Namely, it may be characterized also by magnetic (in particular, antiferromagnetic) or charge order, or even more exotic type of ordering (e.g. staggered flux phase (SF), stripe order [54, 73, 74, 75, 199] or valence-bond solid [78]) Operator $\hat{P}_C$ introduces many-body effects (correlations), hence it is termed the *correlator*. States of the form (7.139) represent thus a natural generalization of (7.137) [33]. They will be called *correlated states* or *correlated wave functions*.

Each choice of either $|\Psi_0\rangle$ or $\hat{P}_C$ defines different $|\Psi\rangle$. One possible form of $|\Psi\rangle$ is the variational state originally proposed by Gutzwiller [20] for the Hubbard model (6.132), i.e.,

$$|\Psi\rangle = \hat{P}_G(\alpha)|\Psi_0\rangle = \prod_i \Big(1 - (1-\alpha)\hat{n}_{i\uparrow}\hat{n}_{i\downarrow}\Big)|\mathrm{FS}\rangle. \tag{7.140}$$

In above, $|\Psi_0\rangle = |\mathrm{FS}\rangle$ is the Fermi-sea wave function. Contrary to $\hat{P} = \hat{P}_G(0)$ appearing in (7.137), $\hat{P}_G(\alpha)$ projects doubly occupied configurations only partially, to a degree dependent on the value of the variational parameter α.

Another important class of correlated trial states $|\Psi\rangle$, generalizing (7.137) are those with $\hat{P}_C$ being in the form of a product of various Jastrow correlators [52, 53, 181], cf. also Eqs. (25)-(28) of Ref. [33]. Here, we are mainly interested in the generalization of (7.137) proposed by Fukushima [76]

$$|\Psi\rangle = \hat{P}^{(F)}(\lambda^F_{i\uparrow}, \lambda^F_{i\downarrow})|\Psi_0\rangle = \prod_i (\lambda^F_{i\uparrow})^{\frac{\hat{n}_{i\uparrow}}{2}} (\lambda^F_{i\downarrow})^{\frac{\hat{n}_{i\downarrow}}{2}} (1 - \hat{n}_{i\uparrow}\hat{n}_{i\downarrow})|\Psi_0\rangle. \tag{7.141}$$

In above, the Gutzwiller projector (6.128) has been replaced by correlator $\hat{P}_C = \hat{P}^{(F)}$, differing from $\hat{P}$ by the presence of the so-called fugacity factors[46] $\lambda^F_{i\sigma}$ [76]. For $\lambda^F_{i\sigma} = 1$, $\hat{P}^{(F)} = \hat{P}$ and then $|\Psi\rangle$ (7.141) reduces to (7.137) if only $|\Psi_0\rangle = |BCS\rangle$. However, within the formalism of Ref. [76], $\lambda^F_{i\sigma}$ are determined from the additional conditions, e.g. that the projection does not change the (local) average particle density or particle density and the local magnetic moment (for more details, see Appendix D, Subsection 11.4.

Finally, let us note, that when the Hubbard model is a starting point of analysis, apart from partially projected states (7.140), $|\tilde{\Psi}\rangle$ of the form

$$|\tilde{\Psi}\rangle = \exp(iS)|\Psi\rangle = \exp(iS)\left(\prod_i (1 - \hat{n}_{i\uparrow}\hat{n}_{i\downarrow})\right)|\mathrm{BCS}\rangle, \tag{7.142}$$

is sometimes used instead $|\Psi\rangle$ (7.137), [48, 61]. Doubly occupied sites in $|\tilde{\Psi}\rangle$ are perturbatively reintroduced by means of the unitary operator $\exp(iS)$ appearing in canonical transformation (6.133).

Here, because our point of departure is the t-J model, we study only the fully projected states (7.141). However, note that within the formalism of Ref. [76], extension to the general case with nonzero double occupancies is also possible, leading to the correlated state of the form

$$|\Psi\rangle = \hat{P}^{(F)}_G(\alpha, \lambda^F_{i\uparrow}, \lambda^F_{i\downarrow})|\Psi_0\rangle = \prod_i (\lambda^F_{i\uparrow})^{\hat{n}_{i\uparrow}} (\lambda^F_{i\downarrow})^{\hat{n}_{i\downarrow}} (1 - (1-\alpha)\hat{n}_{i\uparrow}\hat{n}_{i\downarrow})|\Psi_0\rangle, \tag{7.143}$$

which combines features of (7.140) and (7.141). $|\Psi\rangle$ of the form (7.143) allows to study both the Hubbard and the t-J-U model 6.136.

7.2 Mean-field treatment of Gutzwiller projected state

As mentioned in Introduction, the RVB concept may be implemented either by means of the Variational Monte-Carlo (VMC) methods or by using an appropriate form of mean-field approach. The latter way is partly motivated by the fact, that expectation values with respect to $|\Psi_0\rangle$, and therefore also with respect to $|\Psi\rangle$ (7.139) may be evaluated, at least in principle, using Wick's theorem [136, 165]. In such case, $|\Psi_0\rangle$ is an eigenstate of the effective single-particle RMFT Hamiltonian $\hat{H}_R$, and the latter is obtained from the mean-field treatment of (6.126), cf. e.g. [33, 58, 63, 66]. This point of view is also taken up here.

The crucial point in the construction of RMFT is an analytic evaluation of expectation value of arbitrary operator $\hat{\mathcal{O}}$ with respect to the correlated variational state $|\Psi\rangle$ (7.139), i.e.,

$$\langle\hat{\mathcal{O}}\rangle_C \equiv \frac{\langle\Psi|\hat{\mathcal{O}}|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \frac{\langle\Psi_0|\hat{P}_C\hat{\mathcal{O}}\hat{P}_C|\Psi_0\rangle}{\langle\Psi_0|\hat{P}^2_C|\Psi_0\rangle} \equiv \frac{\langle\hat{P}_C\hat{\mathcal{O}}\hat{P}_C\rangle}{\langle\hat{P}^2_C\rangle}. \tag{7.144}$$

[46] $\lambda^F_{i\sigma}$ in (7.141) should not be confused with the Lagrange multipliers appearing within our method.

Both $\hat{H}_R$ and $|\Psi_0\rangle$ will be subsequently constructed, but are not required at this point. Here, it is only important, that $\hat{H}_R$ is bilinear in fermion creation and/or annihilation operators. In principle, this feature allows us, by applying Wick's theorem, to express the correlated averages $\langle \ldots \rangle_C$ appearing in (7.146) through the uncorrelated ones, $\langle \ldots \rangle$. However, usually the latter step cannot be performed exactly, due to large number (M_W) of all possible contractions, i.e., different uncorrelated averages, appearing in the full Wick's decomposition of $\langle \hat{\mathcal{O}} \rangle_C$ (7.144). Consequently, we evaluate $\langle \hat{\mathcal{O}} \rangle_C$ in an approximate way. First, we have to choose M relevant single-particle operators $\{\hat{A}_s\} = \hat{A}_1, \ldots, \hat{A}_M$ ($M \ll M_W$), the uncorrelated expectation values (mean fields) of which, i.e.,

$$\langle \hat{A}_1 \rangle \equiv A_1, \ldots, \langle \hat{A}_M \rangle \equiv A_M, \qquad \langle \hat{A}_s \rangle \equiv \langle \Psi_0 | \hat{A}_s | \Psi_0 \rangle, \tag{7.145}$$

are assumed to have a non-zero value. We may now rewrite (7.144) as

$$\langle \hat{\mathcal{O}} \rangle_C \approx \langle \hat{\mathcal{O}} \rangle_C^{app} \equiv f_{\mathcal{O}}(\vec{A}). \tag{7.146}$$

Note, that even after selecting the set of relevant mean-fields (again denoted as the components of the vector $\vec{A}$), there exist still many different possible forms of $f_{\mathcal{O}}(\vec{A})$, corresponding to different ways of approximate evaluation of $\langle \hat{\mathcal{O}} \rangle_C$. Each prescription of the form (7.146) will be termed the *renormalization scheme* (RS). In some cases, $\langle \hat{\mathcal{O}} \rangle_C^{app}$ (7.146) may be given a more specific form

$$\langle \hat{\mathcal{O}} \rangle_C \approx \langle \hat{\mathcal{O}} \rangle_C^{app} = g^{\mathcal{O}}(\vec{A}) \frac{\langle \Psi_0 | \hat{\mathcal{O}} | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle} \equiv g^{\mathcal{O}}(\vec{A}) \langle \hat{\mathcal{O}} \rangle, \tag{7.147}$$

where $g^{\mathcal{O}}(\vec{A})$ is termed the *renormalization factor* for $\hat{\mathcal{O}}$ operator (cf. Eq. (15) of Ref. [33]).

Many different renormalization schemes has been proposed, most of them have a form of Eq. (7.147). The simplest take into account only the local (site-dependent) mean fields (e.g. local charge density or magnetization), [33, 78]. Within the more complicated ones, inter-site contractions (i.e., the mean-field quantities defined on bonds, e.g. the pairing amplitude) are also included, cf. Refs. [76, 63, 79]. Some renormalization schemes, important from the point of view of the present work, will be listed and briefly discussed in the next Subsection.

Approximate way of computing correlated averages (7.144) or (7.147) is well known under the name of *Gutzwiller approximation* (GA) [20, 33, 59]. It must be noted here, that in most cases GA is an essentially uncontrolled procedure. Namely, as a consequence of approximate evaluation of $\langle \hat{\mathcal{O}} \rangle_C$ for $\hat{\mathcal{O}} = \hat{H}_e$ (e.g. with $\hat{H}_e = \hat{H}_{tJ}$), it is no longer guaranteed, that $f_{\mathcal{O}}(\vec{A})$ (7.146) provides an upper bound for the ground state energy of $\hat{H}_e$. This fact is regarded as the principal weakness of any GA-based approach.

7.2.1 Exact evaluation of correlated averages and rigorous upper bound for exact ground state energy

It should be mentioned at this point, that in some cases it is possible to evaluate the correlated averages (7.144) in an approximation-free manner. Metzner and Vollhardt [205, 206] were able to compute analytically $\langle \hat{\mathcal{O}} \rangle_C$ (7.144) for $|\Psi\rangle$ (7.140), $\hat{\mathcal{O}} = \hat{H}_{tU}$ (6.132), and $D_S = 1$, i.e for the Hubbard chain (a detailed discussion and some additional comments to the original papers are provided also in Ref. [207]). Later, Gebhard [99, 100] developed further the ideas of Metzner and Vollhardt and proposed a systematic and controlled way of evaluating correlated averages by means of the so-called $1/D_S$ expansion. Similar approach is being currently developed by Bünemann, Schickling and Gebhard [208].

Let us also note, that it is much more difficult (or even impossible) to evaluate the numerator (or denominator) of (7.144) separately, than to compute the whole correlated average. This is

because only in the latter case, the linked cluster theorem may be applied, which makes the task executable.

If (7.144) with $\hat{\mathcal{O}} = \hat{H}_e$ is computed with no approximations, $f_{\mathcal{O}}(\vec{A})$ provides an exact upper bound for the ground state energy of $\hat{H}_e$. Then, $(A_1, A_2, \ldots, A_M) \equiv \vec{A}$ may be treated as variational parameters, with respect to which the r.h.s of (7.146) is being maximized, under the proviso, that the self-consistency constraints are imposed. From the conditions (7.145) it should become evident, that the presence of such constraints is necessary.

7.2.2 Projected versus unprojected quantities

Any approach based on the correlated trial wave functions (7.139) has an unavoidable dual character, in the following sense. Namely, for each physical quantity $\mathcal{O}$, represented by an operator $\hat{\mathcal{O}}$, a careful distinction should be made between correlated (projected) $\langle\hat{\mathcal{O}}\rangle_C$ (7.144) and uncorrelated (unprojected) $\langle\hat{\mathcal{O}}\rangle$ averages. Usually, the former have physical interpretation of e.g. true long-range superconducting order parameter or correlated average kinetic (band) energy. On the other hand, the unprojected (bare) averages are merely elements of the mathematical formalism. They are the building blocks of the projected averages, they also appear in quasiparticle spectrum of the resulting RMFT Hamiltonian.

7.3 Standard formulation of renormalized mean-field theory

7.3.1 RMFT Hamiltonian

After the particular RS is selected, our next task is to construct an appropriate single-particle Hamiltonian $\hat{H}_R$, which is necessary in order to obtain an explicit form of $|\Psi_0\rangle$. Here we present the main steps of the standard RMFT approach, i.e., with RS that can be given the form (7.147), and with $\lambda^F_{i\sigma} = 1$ in (7.141). Modifications caused by application of the variational approach of Part II will be discussed subsequently.

The standard RMFT scheme may be summarized as follows (cf. Fig. 27 of Ref. [33])

1. We start from $\hat{H}_{tJ}$ (6.126), usually in the simplest form (6.135).

2. Next, as an intermediate step, we introduce Hamiltonian, in which the exact projection has been replaced by renormalization factors g^t_{ij} and g^J_{ij}, but with the kinetic-exchange term kept intact, i.e.,

$$\hat{H}^{(ren)}_{tJ} = \sum_{i,j,\sigma} g^t_{ij} t_{ij} c^\dagger_{i\sigma} c_{j\sigma} + \sum_{\langle ij\rangle} g^J_{ij} J_{ij}\ \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j. \tag{7.148}$$

   In the above formula, renormalization factors g^t_{ij} and g^J_{ij} are given by Eq. (7.147), with $\hat{\mathcal{O}} = (c^\dagger_{i\sigma} c_{j\sigma} + \text{H.c.})$ or $\hat{\mathcal{O}} = \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j$, respectively. Their explicit form depends on the approximation used to obtain r.h.s. of (7.147).

   Obviously, Hamiltonian (7.148) is constructed in such way, that within the approximation (7.147) its expectation value with respect to $|\Psi_0\rangle$ is the same as expectation value of the t-J Hamiltonian (6.135) evaluated with respect to the correlated state $|\Psi\rangle$

$$\langle\Psi_0|\hat{H}^{(ren)}_{tJ}|\Psi_0\rangle \equiv \langle\hat{H}_{tJ}\rangle^{app}_C \approx \langle\hat{H}_{tJ}\rangle_C. \tag{7.149}$$

   Nonetheless, due to the presence of the $\hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j$ terms, $\hat{H}^{(ren)}_{tJ}$ is still intractable, and must be simplified further.

3. The renormalized mean-field t-J Hamiltonian $\hat{H}_R$ is obtained from (7.148) by the Hartree-Fock (HF) -type decoupling (4.107) of the $\hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j$ terms. This involves the particle-hole channel, i.e., the bond-order parameter $\langle c^\dagger_{i\sigma} c_{j\sigma}\rangle$, but it may also involve the particle-particle channel $\langle c_{i\sigma} c_{j\bar{\sigma}}\rangle$, i.e., possibility for a nonzero pairing (superconducting solutions). Finally, in the presence of spin or charge ordering, quantities $\langle \sum_\sigma \sigma c^\dagger_{i\sigma} c_{i\sigma}\rangle \equiv m_i$ and $n - \langle \sum_\sigma c^\dagger_{i\sigma} c_{i\sigma}\rangle \equiv n - n_i$, where n_i is the local electron density, may also acquire nonzero values. In what follows, we focus on the case with neither magnetic nor charge ordering (general case may be treated analogously). Explicitly, exchange interaction reads

$$\hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j = \hat{S}^x_i \hat{S}^x_j + \hat{S}^y_i \hat{S}^y_j + \hat{S}^z_i \hat{S}^z_j = \frac{1}{2}(\hat{S}^+_i \hat{S}^-_j + \hat{S}^-_i \hat{S}^+_j) + \hat{S}^z_i \hat{S}^z_j. \tag{7.150}$$

For the $\hat{S}^\sigma_i \hat{S}^{\bar{\sigma}}_j$ part of $\hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j$ (with $\bar{\sigma} = -\sigma = \pm$), HF decoupling (4.107) leads to

$$\begin{aligned} \hat{S}^\sigma_i \hat{S}^{\bar{\sigma}}_j &= c^\dagger_{i\sigma} c_{i\bar{\sigma}} c^\dagger_{j\bar{\sigma}} c_{j\sigma} \rightarrow \\ &- \langle c^\dagger_{i\sigma} c_{j\sigma}\rangle c^\dagger_{j\bar{\sigma}} c_{i\bar{\sigma}} - \langle c^\dagger_{j\bar{\sigma}} c_{i\bar{\sigma}}\rangle c^\dagger_{i\sigma} c_{j\sigma} + \langle c^\dagger_{i\sigma} c_{j\sigma}\rangle \langle c^\dagger_{j\bar{\sigma}} c_{i\bar{\sigma}}\rangle \\ &- \langle c^\dagger_{i\sigma} c^\dagger_{j\bar{\sigma}}\rangle c_{i\bar{\sigma}} c_{j\sigma} - \langle c_{i\bar{\sigma}} c_{j\sigma}\rangle c^\dagger_{i\sigma} c^\dagger_{j\bar{\sigma}} + \langle c_{i\bar{\sigma}} c_{j\sigma}\rangle \langle c^\dagger_{i\sigma} c^\dagger_{j\bar{\sigma}}\rangle. \end{aligned} \tag{7.151}$$

HF decoupling of the $\hat{S}^z_i \hat{S}^z_j = \frac{1}{4}\big(c^\dagger_{i\uparrow} c_{i\uparrow} - c^\dagger_{i\downarrow} c_{i\downarrow}\big)\big(c^\dagger_{i\uparrow} c_{i\uparrow} - c^\dagger_{j\downarrow} c_{j\downarrow}\big)$ term may be performed analogously. In the above expressions, $c^\dagger_{i\sigma}$ ($c_{j\sigma}$) are ordinary fermion creation (annihilation) operators. As a consequence of approximations made above, $\hat{H}_R$ may be taken in a form [63, 68, 73, 74, 75, 78]

$$\begin{aligned} \hat{H}_R(\vec{A}) &= \sum_{\langle ij\rangle\sigma} \big(t_{ij} g^t_{ij} c^\dagger_{i\sigma} c_{j\sigma} + \text{H.c.}\big) - \frac{3}{4} J_{ij} g^J_{ij} (\chi_{ji} c^\dagger_{i\sigma} c_{j\sigma} + \text{H.c.} - |\chi_{ij}|^2) \\ &- \sum_{\langle ij\rangle\sigma} \frac{3}{4} J_{ij} g^J_{ij} (\Delta_{ij} c^\dagger_{j\sigma} c^\dagger_{i\bar{\sigma}} + \text{H.c.} - |\Delta_{ij}|^2). \end{aligned} \tag{7.152}$$

The mean-fields $(A_1, \ldots A_M) = \vec{A}$ are now respectively, the local electron densities n_i, the hopping amplitudes (bond-order parameter) $\chi_{ij\sigma}$, and the RVB gap parameters, Δ_{ij}, both taken for the nearest neighbors ($\langle ij\rangle$). Explicitly, in the case of no magnetic or charge order, two latter quantities are defined as

$$\chi_{ij\sigma} \equiv \langle c^\dagger_{i\sigma} c_{j\sigma}\rangle, \qquad \Delta_{ij} \equiv \langle c_{i\bar{\sigma}} c_{j\sigma}\rangle = \langle c_{j\bar{\sigma}} c_{i\sigma}\rangle. \tag{7.153}$$

Renormalization factors g^t_{ij} and g^J_{ij} appearing in $\hat{H}_R(\vec{A})$ depend on n_i, but may also depend on $\chi_{ij\sigma}$ and Δ_{ij}.

4. The ground state $|\Psi_0\rangle$ of $\hat{H}_R(\vec{A})$ (7.152) is used next in Eq. (7.147) to obtain the renormalization factors $g^{\mathcal{O}}$ appearing in (7.148). By construction, for $\langle\Psi_0|\Psi_0\rangle = 1$, we have

$$\langle \hat{H}_{tJ}\rangle^{app}_C \equiv \langle \Psi_0 | \hat{H}_R | \Psi_0\rangle. \tag{7.154}$$

The last step completes construction of $\hat{H}_R$. What remains, is the solution of the just constructed MF model; we have to determine the optimal values of the mean-fields and the explicit form of the equilibrium MF density operator. This is discussed in the next Subsection.

7.4 Solving renormalized mean-field theory: application of MaxEnt-based variational approach

With an appropriate single-particle mean-field Hamiltonian $\hat{H}_R$ at our disposal, we are faced with the problem of finding the optimal (equilibrium) values of mean-fields appearing in the model. What is crucial for the present discussion, is that due to the MF treatment of the Gutzwiller projection, the RMFT Hamiltonian is not of the Hartree-Fock form with respect to those of its mean-field variables, which appear in the renormalization factors $g^{\mathcal{O}}$. In such case, the unwary application of the variational procedure is in conflict with the basic requirement of self-consistency (cf. discussion in the Introduction and in the Section 4.3 of Part II). Consequently, RMFT is frequently solved by invoking a non-variational self-consistent approach based on the Bogoliubov-de Gennes (BdG) equations [63, 68, 74, 75].

However, this route, apart from its inherent drawbacks, discussed in detail in Part II, encounters a difficulty in the case of RMFT formalism of Ref. [76], in the following sense. Namely, the renormalization scheme of the latter Reference is not of the form (7.147), but of a more general form (7.146). In general case, this feature does not allow for an unambiguous identification of the renormalization factors and hence for the unique construction of an effective RMFT Hamiltonian within the standard scheme presented above. Obviously, in such a situation, different choices of the renormalization factors yield different RMFT Hamiltonians. Consequently, within the non-variational self-consistent approach we obtain also different BdG equations 4.8), and in effect, different predictions of obtained models.

Another possibility is to treat $\hat{H}_R$ within the self-consistent MF variational approach introduced in the Part II. Interestingly, in such case the MF Hamiltonian is always uniquely defined (cf. Subsection 4.8), hence in particular also for the RS of Ref. [76].

7.4.1 Finite temperature and mixed correlated states

Gutzwiller approximation in a form (7.146) or (7.147) is devised to study the ground state properties of the Hubbard or t-J Hamiltonian at $T = 0$. However, as discussed in Subsection 12.2, it is legitimate study $T = 0$ situation as the $T \to 0$ ($\beta \to \infty$) limit of the $T > 0$ case. In order to apply the present finite-temperature formalism to study RMFT of the t-J model, we have to replace all the pure states appearing in the previous formulas by the corresponding mixed states, i.e.,

$$|\Psi_0\rangle\langle\Psi_0| \to \hat{\rho}_0, \qquad |\Psi\rangle\langle\Psi| \to \hat{\rho}_{\mathrm{C}} = \hat{P}_C \hat{\rho}_0 \hat{P}_C. \tag{7.155}$$

The explicit form of $\hat{\rho}_0$ will be provided below. Similarly to its pure counterpart $|\Psi\rangle$, density operator $\hat{\rho}_{\mathrm{C}}$ is not normalized, i.e., $\mathrm{Tr}[\hat{\rho}_{\mathrm{C}}] \neq 1$.

Note, that we do not claim that the RMFT formalism is valid and physically relevant at arbitrary high temperatures. Here, introduction of the nonzero temperature is a purely formal step, which is necessary in order to use the present approach based on the MaxEnt principle. However, solutions obtained for non-zero, but sufficiently low T are for all practical purposes identical to those of the true $T = 0$ analysis based on the ground state wave function. Therefore, all the results obtained in this Thesis refer in fact to the $T = 0$ situation.

However, let us note, that attempts to generalize the Gutzwiller approximation to finite temperatures have also been made [202, 209, 210, 211]. In particular, in a recent paper [202], a finite-temperature treatment of the t-J-U model (6.136) has been proposed. Within this approach, projected density operator $\hat{\rho}_{\mathrm{C}}$ (7.155) is treated as a proper finite-temperature counterpart $|\Psi\rangle$. Using $\hat{\rho}_{\mathrm{C}}$, both the entropy and the free energy are computed. Some comments on the formalism of Ref. [202] are provided in [83]; a more detailed analysis and possible implementation of TRMFT within our approach is postponed for further studies [212].

7.4.2 Formalism of Part II: application to t-J model

Now we are ready to apply the formalism developed in Part II to the RMFT of the t-J model. To make the subsequent discussion self-contained, we recall the key points of our method.

1. We start from a given initial MF Hamiltonian $\hat{H}(\vec{A}) \equiv \hat{H}$ (4.1), which in the present case is given by $\hat{H}_R$ (7.152). Next, we supplement $\hat{H} = \hat{H}_R$ with the constraint terms (4.13). As mentioned previously, non-Hartree-Fock character of $\hat{H}_R$ is caused by the presence of the renormalization factors g_{ij}^t and g_{ij}^t. Consequently, in order to obtain $\hat{H}_\lambda$ (4.15) corresponding to $\hat{H}_R$ (and from now on denoted by $\hat{H}_{R\lambda}$) we introduce the constraint terms, which involve the mean fields appearing in g_{ij}^t and g_{ij}^J. Therefore, from (7.152) we obtain

$$\begin{aligned}\hat{H}_{R\lambda} = \hat{H}_R \;&-\; \sum_i \lambda_i^{(n)} \big(\sum_\sigma c_{i\sigma}^\dagger c_{i\sigma} - n_i \big) - \sum_{\langle ij \rangle \sigma} \big(\lambda_{ij}^{(\chi)} (c_{i\sigma}^\dagger c_{j\sigma} - \chi_{ij}) + \text{H.c.} \big) \\ &-\; \sum_{\langle ij \rangle \sigma} \big(\lambda_{ij}^{(\Delta)} (c_{i-\sigma} c_{j\sigma} - \Delta_{ij}) + \text{H.c.} \big). \end{aligned} \tag{7.156}$$

   In above, 'H.c.' stands for 'Hermitian conjugate'.

2. As a next step, MF density operator $\hat{\rho}_\lambda \equiv \hat{\rho}_{R\lambda}$ is constructed (cf. Eq. (4.29)). We have

$$\hat{\rho}_{R\lambda} = \mathcal{Z}_{R\lambda}^{-1} e^{-\beta(\hat{H}_{R\lambda} - \mu \hat{N})}, \quad \mathcal{Z}_{R\lambda} = \text{Tr}[e^{-\beta(\hat{H}_{R\lambda} - \mu \hat{N})}]. \tag{7.157}$$

   $\hat{\rho}_{R\lambda}$ will play the role of $\hat{\rho}_0$ appearing in (7.155).

3. We define generalized Landau grand potential (cf. Eqs. (4.33)), i.e.,

$$\mathcal{F}(\vec{A}, \vec{\lambda}) \equiv -\beta^{-1} \ln \mathcal{Z}_{R\lambda}(\vec{A}, \vec{\lambda}). \tag{7.158}$$

   Next, $\mathcal{F}(\vec{A}, \vec{\lambda})$ is used to obtain the optimal values of $\vec{A}$, and $\vec{\lambda}$, i.e., $\vec{A} = \vec{A}_0$, $\vec{\lambda} = \vec{\lambda}_0$. There are two possibilities:

   (a) Within the present variational (*var*) approach, $\vec{A}_0$ and $\vec{\lambda}_0$ are given by those solution of Eqs. (4.34), i.e.,

$$\nabla_A \mathcal{F} = \vec{0}_M, \quad \nabla_\lambda \mathcal{F} = \vec{0}_M, \tag{7.159}$$

   for which $\mathcal{F}(\vec{A}, \vec{\lambda})$ has the lowest value.

   (b) By taking the derivatives with respect to $\vec{\lambda}$ only, and subsequently putting $\vec{\lambda} = \vec{0}$,

$$\nabla_\lambda \mathcal{F} = \vec{0}_M, \quad \vec{\lambda}_0 = \vec{0}_M, \tag{7.160}$$

   we obtain Bogoliubov-de Genes (BdG) self-consistent (s-c) equations (cf. Eqs. 4.38). In such case, we denote the optimal values of mean-fields by $\vec{A}_{sc}$, and the chemical potential by $\tilde{\mu} = \mu_{sc}$. The latter quantity corresponds to $\tilde{\mu} = \mu + \lambda$ within the *var* method.

4. Thermodynamic grand potential $\Omega(T, V, \mu)$ (4.47) and the free energy $F(T, V, N)$ (4.58) are defined respectively as

   (a) Within the *var* method we have (cf. Eqs. 4.47 and 4.58)

$$\Omega(T, V, \mu) = \mathcal{F}(T, V, \mu; \vec{A}_0(T, V, \mu), \vec{\lambda}_0(T, V, \mu)), \qquad F = \Omega + \mu N. \tag{7.161}$$

(b) For the non-variational *s-c* approach, we have

$$\Omega_{sc}(T,V,\mu_{sc}) = \mathcal{F}(T,V,\mu;\vec{A}_{sc}(T,V,\mu),\vec{0}), \qquad F_{sc} = \Omega_{sc} + \mu_{sc}N. \tag{7.162}$$

Because we work with fixed total particle number N, we mainly use F, and not Ω.

If the Landau potential $\mathcal{F}_z^{(\Lambda)}(T,V,\mu;\vec{A})$ (4.44) is required, it may be constructed, at least in principle, according to the prescription given in Section 4.5. This task usually have to be accomplished by means of the numerical analysis. We will not use $\mathcal{F}_z^{(\Lambda)}$ in the subsequent presentation of the numerical results, but let us note that plots of Landau potential may help to distinguish the character of the phase transition. This may be difficult from the sole analysis of the behavior of the order parameter near the transition point. Also, variational parameters of the non-MF character ($\vec{b}$) will not be used here, consequently we have omitted $\vec{b}$ in Eqs. (7.161) - (7.156)

Even without a numerical analysis, it can be shown that for most of the choices of g_{ij}^t and g_{ij}^t, at least one of the Lagrange multipliers appearing in (7.156) has nonzero equilibrium value, i.e., $\vec{\lambda}_0 \neq \vec{0}$ [113]. Therefore, from the very beginning we may be sure the present variational approach yields results that differ from those of the non variational *s-c* treatment.

Summarizing, we modify the standard RMFT scheme as follows. Instead of $\hat{H}_R$ (7.152), we use $\hat{H}_{R\lambda}$ (7.156). In (7.144) we replace pure by mixed states according to (7.155), with $\hat{\rho}_0 = \hat{\rho}_{R\lambda}$. Next, the generalized Landau potential $\mathcal{F}$ (7.158) is constructed, and subsequently used to obtain Eqs. (7.159). The solution of the latter equations, characterized by the lowest value of $\mathcal{F}$, provides the optimal set of the mean-fields.

The above scheme may be regarded as a both natural and necessary extension of the standard RMFT. It allows for the consistent application of the variational method based on the maximum entropy principle. Also, it allows us to implement both variational (*var*, Eqs. 7.159) and non-variational (*s-c*) approach based on BdG equations, Eqs. (7.160) within a single framework. Therefore, this particular formulation is convenient for a comparison of the *var* and *s-c* methods. This composes our main task in the next Section.

7.4.3 Alternative formulation of renormalized mean-field theory

By making use of the results of Section 4.8, we may formulate RMFT of the *t-J* model in an alternative way, which is simpler, more transparent and in many cases better adaptable for applications. First, let us note that the grand Hamiltonian $\hat{H}_R - \mu\hat{N}$ corresponding to $\hat{H}_R$ (7.152) is of the form (4.78). Therefore, we may apply transformations (4.77) to $\hat{H}_R$ and obtain new MF Hamiltonian $\hat{H}_R^{(\sim)}$, according to

$$\hat{H}_R^{(\sim)} \equiv \langle \hat{H}_R \rangle \hat{\mathbf{1}}_{D_H} = \langle \hat{H}_e \rangle_C^{app} \hat{\mathbf{1}}_{D_H} \equiv W(\chi_{ij\sigma}, \Delta_{ij}, n_{i\sigma}) \hat{\mathbf{1}}_{D_H}. \tag{7.163}$$

In above, $\hat{\mathbf{1}}_{D_H}$ denotes a unit matrix, which will be omitted for simplicity in the following formulas. Thus, $\hat{H}_\lambda^{(\sim)}$ (4.81) (denoted from now on as $\equiv \hat{H}_{R\lambda}^{(\sim)}$), is given by

$$\begin{aligned} \hat{H}_{R\lambda}^{(\sim)} &= -\sum_{\langle ij\rangle\sigma} \Big(\tilde{\eta}_{ij\sigma} \big(c_{i\sigma}^\dagger c_{j\sigma} - \chi_{ij\sigma} \big) + \text{H.c.} \Big) - \sum_{i\sigma} \tilde{\lambda}_{i\sigma}^{(n)} \big(\hat{n}_{i\sigma} - n_{i\sigma} \big) \Big) \\ &- \sum_{\langle ij\rangle} \Big(\tilde{\gamma}_{ij} \big(\hat{\Delta}_{ij} - \Delta_{ij} \big) + \text{H.c.} \Big) + W(\chi_{ij\sigma}, \Delta_{ij}, n_{i\sigma}). \end{aligned} \tag{7.164}$$

In above, again we have $n_{i\sigma} = \langle \hat{n}_{i\sigma} \rangle$, $\chi_{ij\sigma} = \langle c_{i\sigma}^\dagger c_{j\sigma} \rangle$. However, now the *d*-wave pairing operators $\hat{\Delta}_{ij}$ are defined in a symmetric form (cf. Eq. (6.130)), from which their singlet character is explicit

$$\hat{\Delta}_{ij} = \frac{1}{2} \left(c_{i\uparrow} c_{j\downarrow} - c_{i\downarrow} c_{j\uparrow} \right). \tag{7.165}$$

Consequently, superconducting gap parameter is defined as $\Delta_{ij} = \langle \hat{\Delta}_{ij} \rangle$, but this difference with the previously given one is inessential. The averages $\langle \ldots \rangle$ in (7.164) are defined with the help of density operator $\hat{\rho}_\lambda = \hat{\rho}_0$, and not by the ground state $|\Psi_0\rangle$. Note, that a nontrivial (i.e., not proportional to $\hat{\mathbf{1}}_{D_H}$) operator part is provided only by the constraint terms.

Having $\hat{H}^{(\sim)}_{R\lambda}$ (7.164) at our disposal, we may easily proceed along the lines of analysis given in Section 4.8. We can simply use $\hat{H}^{(\sim)}_{R\lambda}$ instead of $\hat{H}_{R\lambda}$ within the standard scheme supplemented with the MaxEnt variational method (see the previous Subsection).

Formulation of the RMFT of the t-J model presented above may be generalized to other lattice-fermion models, defined by an arbitrary non-MF 'exact' many-body Hamiltonian $\hat{H}_e$, and to arbitrary trial correlated state $\hat{\rho}_C$ (7.155). Therefore, for the sake of completeness and generality, we summarize below the main steps of this prescription without referring to any particular Hamiltonian or variational state. Later, this formulation will be used for $\hat{H}_e = \hat{H}_{tJ}$ and $\hat{\rho}_C$ (7.155), with $\hat{P}_C = \hat{P}^{(F)}_C$ appearing in (7.141). The steps are

1. We begin with a given non-MF Hamiltonian $\hat{H}_e$ and the correlator $\hat{P}_C$. Next, we anticipate the existence of (not yet specified) MF single-particle density operator $\hat{\rho}_0$ (7.155); this allows us to apply Wick's theorem [136, 165] in order to compute the following quantity

$$W^{(e)} \equiv \langle \hat{H}_e \rangle_C \equiv \frac{\mathrm{Tr}[\hat{\rho}_C \hat{H}_e]}{\mathrm{Tr}[\hat{\rho}_C]} = \frac{\mathrm{Tr}[\hat{P}_C \hat{\rho}_0 \hat{P}_C \hat{H}_e]}{\mathrm{Tr}[\hat{P}_C \hat{\rho}_0 \hat{P}_C]} = \frac{\mathrm{Tr}[\hat{\rho}_0 \hat{P}_C \hat{H}_e \hat{P}_C]}{\mathrm{Tr}[\hat{P}_C \hat{\rho}_0 \hat{P}_C]} = \frac{\langle \hat{P}_C \hat{H}_e \hat{P}_C \rangle}{\langle \hat{P}^2_C \rangle}. \quad (7.166)$$

   Here, from the very the beginning we use mixed states, $\hat{\rho}_0$ and $\hat{\rho}_C$ (7.155), instead of their pure correspondents $|\Psi_0\rangle$ and $|\Psi\rangle$. As discussed in Subsection 7.2, due to the large number (M_W) of the possible contractions, i.e., uncorrelated averages of the operators bilinear in fermion creation and/or annihilation operators, $\langle \hat{H}_e \rangle_C$ (7.166) usually have to be computed in an approximate manner. When doing this, we select the relevant single-particle operators $\{\hat{A}_s\} = \hat{A}_1, \ldots, \hat{A}_M$, and the corresponding mean-fields, $A_1 \equiv \langle \hat{A}_1 \rangle, A_2 \equiv \langle \hat{A}_2 \rangle, \ldots, A_M \equiv \langle \hat{A}_M \rangle$, $(A_1, A_2, \ldots, A_M) \equiv \vec{A}$, $M \ll M_W$. In effect, $W^{(e)}$ is approximated by $W(\vec{A})$,

$$W^{(e)} = \langle \hat{H}_e \rangle_C \approx \langle \hat{H}_e \rangle^{app}_C \equiv W(\vec{A}), \quad (7.167)$$

   which is a function of the selected mean-fields. Again, for given $\hat{H}_e$, different choices of $\hat{P}_C$, as well as various ways of evaluating $W(\vec{A})$ yield different RS (7.146) or (7.147).

2. Next, $W(\vec{A})$ is supplemented with an operator constraint term for each $\hat{A}_s \in \{\hat{A}_1, \ldots, \hat{A}_M\}$. First, this step allows to treat each A_s as variational parameters. Second, it yields the MF Hamiltonian of the form (4.81), and the corresponding grand Hamiltonian

$$\hat{K}^{(\sim)}_{R\lambda}(\vec{A}, \vec{\lambda}) = \hat{H}^{(\sim)}_{R\lambda}(\vec{A}, \vec{\lambda}) - \mu \hat{N} = \sum_{s=1}^{M} \lambda_s (\hat{A}_s - A_s) + W(\vec{A}) - \mu \hat{N}. \quad (7.168)$$

   Note, that the $\mu \hat{N}$ term may be included in (7.166) as well, and then its expectation value appears in $W(\vec{A})$, but is no longer present on the operator level in (7.168). This is a special case of transformations (4.79), which lead to the equivalent MF Hamiltonian, differing only with respect to the equilibrium value of the Lagrange multiplier λ_1, coupled to $\hat{A}_1 = \hat{N}$.

3. Making use of $\hat{K}^{(\sim)}_{R\lambda}$ (7.168), we construct single-particle density operator $\hat{\rho}_\lambda$ (4.29) in a standard manner, i.e.,

$$\hat{\rho}_{R\lambda} = \mathcal{Z}^{-1}_{R\lambda} e^{-\beta \hat{K}^{(\sim)}_{R\lambda}}. \quad (7.169)$$

   $\hat{\rho}_{R\lambda}$ plays the role of $\hat{\rho}_0$ in (7.166). This step closes the whole scheme.

Please note, that if $\langle \hat{H}_e \rangle_C^{app} \equiv W(\vec{A})$ is computed exactly, it may be regarded as an exact upper bound for the ground-state energy also in the present case with $T > 0$. This is because at sufficiently low T, the entropic part of the free energy is very small, and there is practically no difference between the ground state energy and the free energy.

The above formulation is particularly convenient in case of the formalism of Ref. [76], which consist certain prescriptions for the evaluation of (7.166) with $\hat{H}_e = \hat{H}_{tJ}$. However, in Ref. [76] a problem of construction of the single-particle effective Hamiltonian was only outlined, though not fully accomplished. It has been mentioned, that in order to construct and solve RMFT based on RS of [76], one may utilize formalism of Ref. [71]. Relation of the latter formalism to the present treatment has been discussed in Subsection 4.8.

Also, as pointed out in Section 4.8, the advantage of the RMFT formulation given by Eqs. (7.166)-(7.169) is that one can solve half of Eqs. (7.159), ($\nabla_A \mathcal{F} = 0$) analytically, and eliminate Lagrange multipliers in favor of the mean-fields. Only the latter half of Eqs (7.159), i.e., $\nabla_\lambda \mathcal{F} = 0$ must be solved numerically for the mean-fields. However, this particular formulation cannot be used to obtain the non-variational self-consistent BdG equations. Therefore, it is applied only in Section 9, after the comparison of the numerical results for both variational and the non-variational approach is made in Section 8.

7.4.4 Choice of relevant mean-fields

So far, we have made no restriction on the choice of mean-fields $\langle \hat{A}_s \rangle = A_s$, $s = 1, 2, \ldots, M$ relevant for the problem at hand. However, as mentioned in Section 7.2, within any approximate method of computing correlated average values (7.144), the number of mean fields (M), appearing in $W(\vec{A})$ is usually much smaller than M_W, the number of all possible different Wick's contractions, i.e., averages of fermionic bilinears. If we assume *a priori* that some mean-field averages do not appear in $W(\vec{A})$, i.e., $A_t = 0$ for $t = M+1, \ldots, M_W$, we may encounter two situations.

First, after the MF model is solved, it may turn out that indeed $A_{t0} = 0$, i.e., that the equilibrium value of A_t vanishes. This situation may correspond to some unbroken symmetry present in our model, e.g. the translational symmetry ($\forall i, j: \ i \equiv j$), or to absence of the superconducting order ($\forall i, j: \ \Delta_{ij} = 0$). As another examples of such situation, we may invoke lack of magnetic order ($\forall i: \ m_i = 0$), or the vanishing expectation value of a 'spin flip' terms, i.e., $S_{ij}^{\sigma} \equiv \langle c_{i\sigma}^{\dagger} c_{j\bar{\sigma}} \rangle = 0$, either with $i = j$ or $i \neq j$ (this assumption corresponds to an unique, i.e., site independent choice of the quantization axis).

However, as already mentioned in Subsection 4.4.5, it may happen, that even if certain A_s does not appear in $W(\vec{A})$ (i.e., we put $A_s = 0$ 'by hand' when computing $W(\vec{A})$), we may still obtain nonzero equilibrium value of such mean field, $A_{s0} \equiv \langle \hat{A}_s \rangle_0 \neq 0$. As an example, let us consider average hopping of electrons with a given spin σ to the nearest ($\langle ij \rangle$) and next-nearest ($\langle\langle ij \rangle\rangle$) neighbors, $\chi^{(1)} \equiv \sum_{\langle ij \rangle} \langle c_{i\sigma}^{\dagger} c_{j\sigma} \rangle$ and $\chi^{(3)} \equiv \sum_{\langle\langle ij \rangle\rangle} \langle c_{i\sigma}^{\dagger} c_{j\sigma} \rangle$, respectively. On a square lattice, assuming translational invariance of the mean-field solution, we have $\chi^{(m)} = \langle \sum_{\mathbf{k}} \epsilon^{(m)} c_{\mathbf{k}\sigma}^{\dagger} c_{\mathbf{k}\sigma} \rangle$ with $m = 1, 3$; $\epsilon^{(1)} = 2(\cos k_x + \cos k_y)$ and $\epsilon^{(3)} = 4(\cos k_x \cos k_y)$. It should be clear, that in general we should have $\chi^{(1)} \neq 0$ and within any RMFT approach to the t-J model this mean-field is always present. On the other hand, in order to simplify $W(\vec{A})$, one may assume $\chi^{(3)} = 0$ from the very beginning. Indeed, $\chi^{(3)}$ is rarely present as an independent mean-field variable in RMFT Hamiltonians. Nonetheless, still, for an arbitrary filling n we usually obtain $\chi_0^{(3)} \neq 0$. This property of the model follows solely from the geometry of the Fermi surface and the particular form of the dispersion relation. Similar remark should be valid for the hopping or pairing amplitude between further neighbors, but also for the local pairing amplitude, $\Delta_{ii}^{*} = \langle c_{i\uparrow}^{\dagger} c_{i\downarrow}^{\dagger} \rangle$.

7.5 Renormalization schemes used in the present work

7.5.1 The simplest from of renormalization factors

In the simplest case, renormalization factors (7.147)) for the kinetic energy (g^t_{ij}) and the spin-exchange interaction (g^J_{ij}) depend only on the local hole densities $x_i \equiv 1 - n_i$, cf. [33, 58, 64, 65, 67, 68, 70, 72, 73, 78],

$$g^t_{ij} = \sqrt{\frac{4x_i x_j}{(x_i+1)(x_j+1)}}, \qquad g^J_{ij} = \frac{4}{(x_i+1)(x_j+1)}. \tag{7.170}$$

In the framework of our method, the constraint terms of the following form

$$\hat{Q}^{(n)}_i = -\lambda^{(n)}_{i\sigma}\big((\hat{n}_{i\uparrow} + \hat{n}_{i\downarrow}) - n_i\big) \tag{7.171}$$

should be added to $\hat{H}_R(\vec{A})$ (7.152) in order to obtain $\hat{H}_{R\lambda}$ (7.156) (cf. also Ref. [70]). However Lagrange multiplier $\lambda^{(n)} = \lambda_1$ coupled to the total particle number, and distinct from the chemical potential μ should be introduced even in the homogeneous case ($n = n_i$). Note, that this is a generic feature of MF models with n-dependent MF Hamiltonian, as discussed in Subsection 4.7, and in the context of RMFT it has been noticed e.g. in Refs. [58, 70, 85]. Still, for g^t and g^J (7.170), and in the case of nonmagnetic, homogeneous states, both *var* and *s-c* methods leads to the same equilibrium values of mean-fields (see the related discussion in Subsection 4.7).

If magnetic order is present, and renormalization factors depend also on the local magnetization,[47] we have to add the constraints of the form

$$\hat{Q}^{(m)}_i = -\lambda^{(m)}_i\big((\hat{n}_{i\uparrow} - \hat{n}_{i\downarrow}) - (n_{i\uparrow} - n_{i\downarrow})\big) \quad \text{or} \quad \sum_\sigma \hat{Q}^{(n)}_{i\sigma} = -\sum_\sigma \lambda^{(n)}_{i\sigma}(\hat{n}_{i\sigma} - n_{i\sigma}), \tag{7.172}$$

depending on the choice of independent MF variables (n_i and $m_i = n_{i\uparrow} - n_{i\downarrow}$ or $n_{i\uparrow}$ and $n_{i\downarrow}$). Also, from Eq. (4.30), with A_w equal χ_{ij} or Δ_{ij}, it follows, that for RS (7.170) we have $\lambda^{(\chi)}_{ij0} = \lambda^{(\Delta)}_{ij0} = 0$. This is because neither χ_{ij} nor Δ_{ij} appear in g^t_{ij} and g^J_{ij}.

If, apart form n_i (and possibly m_i), g^t and g^J depend also on the inter-site mean-field variables ([63, 66, 74, 75, 76]), the MF Hamiltonian $\hat{H}_R$ has to be modified further. Namely, the constraint terms $\lambda^{(\chi)}_{ij}(c^\dagger_{i\sigma}c_{j\sigma} - \chi_{ij})$ and $\lambda^{(\Delta)}_{ij}(c_{i-\sigma}c_{j\sigma} - \Delta_{ij})$ have to be added to $\hat{H}_R$ (7.152) in such case (cf. Eqs. (7.156)), and after the model is solved, we obtain $\lambda^{(\chi)}_{ij0} \neq 0$, $\lambda^{(\Delta)}_{ij0} \neq 0$.

7.5.2 Renormalization scheme of Fukushima

As pointed out previously, Gutzwiller approximation used Ref. [76] does not reduce to a simple multiplication of operators by the corresponding renormalization factors (7.147), but is provided in a more general form (7.146). For example, for the hopping amplitude for spin-up electrons we have (Eq. (15) of Ref. [76], but here written using our own notation), i.e.,

$$\langle c^\dagger_{i\uparrow}c_{j\uparrow}\rangle_C \approx \sqrt{\frac{1-n_i}{1-n_{i\uparrow}}}\sqrt{\frac{1-n_j}{1-n_{j\uparrow}}}\left(\chi_{ij\uparrow} - \chi_{ij\downarrow}\frac{\chi_{ij\uparrow}\chi^*_{ij\downarrow} + \Delta_{ji}\Delta^*_{ij}}{(1-n_{i\downarrow})(1-n_{j\downarrow})}\right), \tag{7.173}$$

The above form of $\langle c^\dagger_{i\uparrow}c_{j\uparrow}\rangle_C$ does not allow for a straightforward identification of the renormalization factor g^t. However, we want to compare the variational and non-variational approaches

[47] We assume, that there exist a unique (global) quantization axis, say z axis. The case of non-collinear magnetic ordering leads to further modifications.

also on the example of RMFT based on Ref. [76]. Making use of (7.173) for a simple non-magnetic states ($n_{i\uparrow} = n_{i\downarrow} = n_i/2$, $\langle c^{\dagger}_{i\uparrow} c_{j\uparrow}\rangle = \chi_{ij\uparrow} = \chi_{ij\downarrow} \equiv \chi_{ij}$) and real superconducting pairing amplitude of a singlet character, $\Delta_{ij} = \Delta_{ji} = \Delta^{*}_{ji}$, we may define g^t and g^J as

$$g^t_{(I)ij} = \sqrt{\frac{2(1-n_i)}{2-n_i}}\sqrt{\frac{2(1-n_j)}{2-n_j}}\left(1 - 4\frac{\chi^2_{ij} + \Delta^2_{ij}}{(2-n_i)(2-n_j)}\right), \tag{7.174}$$

$$g^J_{(I)ij} = \frac{4}{(2-n_i)(2-n_j)}. \tag{7.175}$$

The prescription given by Eqs. (7.174) and (7.175) will be referred to as a RS (I). Further comments on the formalism of Ref. [76] are provided in Appendix D (Subsection 11.4). Note, that within our method, and in contrast to the case of the *s-c* approach, different possible choices of g^t_{ij} and g^J_{ij} are equivalent, if only they lead the same functional form of $W(\chi_{ij\sigma}, \Delta_{ij}, n_{i\sigma})$ (7.163).

7.5.3 Renormalization scheme of Sigrist et al.

Renormalization scheme (I) will be compared with that based on renormalization factors taken from Ref. [63], namely

$$g^t_{(II)ij} = \sqrt{\frac{4x_i x_j (1-x_i)(1-x_j)}{(1-x_i^2)(1-x_j^2) + 8(1-x_i x_j)|\chi_{ij}|^2 + 16|\chi_{ij}|^4}}, \tag{7.176}$$

$$g^J_{(II)ij} = \frac{4(1-x_i)(1-x_j)}{(1-x_i^2)(1-x_j^2) + 8x_i x_j(|\Delta_{ij}|^2 - |\chi_{ij}|^2) + 16(|\Delta_{ij}|^4 + |\chi_{ij}|^4)}, \tag{7.177}$$

and referred to as RS (II). This RS has been used, among others, in Refs. [63, 74] and [75] and also in our previous work, [82].

7.5.4 Renormalization scheme of Ogata and Himeda

We want to comment also on the RS proposed by Ogata and Himeda [66]. RMFT based on this RS has been used to study the coexistence of the d-wave superconductivity and antiferromagnetism, and yields the results similar to those of VMC approach. However, this particular RS has also a serious drawback, as explained below. In the absence of antiferromagnetic order ($m = 0$), renormalization factors g^t and g^J (Eqs. (2.12)-(2.16) of Ref. [66]) take the form

$$g^t_{(III)} = \frac{2(1-n)}{2-n} \cdot \frac{(2-n)^2 - 4(\Delta^2 + \chi^2)}{(2-n)^2}, \tag{7.178}$$

$$g^J_{(III)xy} = g^J_z = g^J = \frac{4}{(2-n)^2} \cdot a^{-7}(n, \chi, \Delta; p), \tag{7.179}$$

where

$$a(n, \chi, \Delta; p) = 1 + \frac{8(1-n)^2(\Delta^2 - \chi^2) + \frac{16}{p}(\Delta^2 + \chi^2)^2}{n^2(2-n)^2}. \tag{7.180}$$

In the above equation, $p = 1$ corresponds to the original version of the formalism (cf. Eqs. 2.14 of Ref. [66]), whereas $p = 2$ refers to what Authors call a 'slight adjustment' (Eqs. 2.15). Those two choices will be denoted here as IIIa and IIIb, respectively. Please note, that even before any numerical analysis is carried out, it seems very likely that such change seriously modifies the final results. Namely, although the factor $1/p$ multiplies an expression $16(\Delta^2 + \chi^2)^2$, which is

expected to be of order of at most 10^{-3}, and the difference between $a(n,\chi,\Delta;1)$ and $a(n,\chi,\Delta;2)$ should be of the same order, the exponent 7 in denominator of $g^J_{(III)}$ may lead to much larger difference between values of $g^J_{(III)}$ for $p=1$ and $p=2$. Indeed, in Subsection 8.1 it will be shown that the differences between those two cases are rather pronounced.

Moreover, for us it seems that this 'slight adjustment' is made *ad hoc*, and its only justification is a better consistency of such obtained results with the VMC results. Surely, an agreement of any RMFT approach with the VMC method should be considered as an advantage. Nonetheless, in our opinion the arbitrary change $p=1 \rightarrow p=2$ in Eq. (7.180)) lowers value of the formalism of Ref. [66]. Its another disadvantage is that it has been devised only for the simplest version of the t-J model (6.135) and for the homogeneous solutions with $\Delta=\Delta^*$, $\chi=\chi^*$. We will analyze the formalism of Ref. [66] very briefly, and only in order to show the relatively large differences between IIIa and IIIb. In will not be used to obtain any new results.[48]

7.5.5 Renormalized superconducting order parameter

Within RMFT, it is not the bare average $\Delta_{ij} \equiv \langle \hat{\Delta}_{ij} \rangle$ but rather its correlated counterpart, $\Delta_{Cij} \equiv \langle \hat{\Delta}_{ij} \rangle_C$, which has an interpretation of the true superconducting order parameter. For any RS of the form (7.147), we have $\Delta_{Cij} = g^{\Delta} \Delta_{ij}$. For the simplest renormalization schemes, we have $g^{\Delta}_{ij} = g^t_{ij}$ [33]. This is also the case for RS (II). On the other hand, within RS (I), for simple homogeneous, non-magnetic states, the corresponding g^{Δ} factor can be defined as (cf. Eq. (18) of Ref. [76])

$$RS(I) \qquad g^{\Delta}_{ij} = \sqrt{\frac{2(1-n_i)}{2-n_i}} \sqrt{\frac{2(1-n_j)}{2-n_j}} \left(1 + \frac{4(\chi_{ij}\chi^*_{ij} + \Delta_{ji}\Delta^*_{ij})}{(2-n_i)(2-n_j)} \right), \tag{7.181}$$

Note, that (7.181) differs from (7.174). Also in Ref. [66], a separate expression for $g^{\Delta} \neq g^t$, is provided (Eq. (6.21) of that Reference).

[48] It should be noted at this point, that the way in which the mean-field model is solved in Ref. [66] does not correspond directly to either what we term *var* method, nor to the non-variational *s-c* treatment. Authors introduce two MF Hamiltonian, one MF single-particle Hamiltonian of the RMFT type, and the second of the Hartree-Fock form. Consequently, two kinds of mean-fields appear, i.e., similarly to our method the number of the mean-field variables is doubled. It seems that at least for this particular example, the method of Ogata and Himeda is equivalent to the $\beta \rightarrow \infty$ limit of our method (*var*). We will not analyzed this issue as inessential from our point of view. To solve the RMFT model with RS of Ref. [66] we use our *var* approach.

8 Results I: Comparison of two methods of approach and different renormalization schemes

In this Section we present the first part of our numerical analysis. Namely, we analyze the simplest form of the t-J Hamiltonian (6.135). For that purpose, we use the standard formulation of RMFT as given by Eq. (7.148)-(7.154). Other forms of the t-J model, including the most complete form (6.126) with $c_1 = c_2 = 1$ are analyzed in Section 9.

In Subsection 8.1 we analyze plain d-wave superconducting state (dSC) with no magnetic order. Next, in Subsection 8.2 we study the staggered-flux (SF) non-superconducting solution [62, 75, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 194]. Finally, Subsection 8.3 is devoted to analysis of the Pomeranchuk instability (PI) [125, 126, 127, 128, 129, 130], i.e., normal state characterized by the broken discrete rotational C_{4v} symmetry.

On the example of those simple MF states we compare the present variational (*var*) method and the non-variational approach based on the BdG self-consistent (*s-c*) equations. Also, we compare different renormalization schemes (RS). It will be shown, that the obtained results depend in a nontrivial way on either different method or different RS.

8.1 Superconducting d-wave (dSC) solution

Solution analyzed in this Subsection possesses full symmetry of the underlying square lattice, only the superconducting order parameter is assumed to have $d_{x^2-y^2}$ symmetry. Consequently, we are left with three independent mean fields $\vec{A} = (n, \chi, \Delta)$; $\chi_x = \chi = \chi_y$, $\Delta_x = \Delta = -\Delta_y$, and with the same number of the corresponding Lagrange multipliers, $\vec{\lambda} = (\lambda, \lambda^\chi, \lambda^\Delta)$, where $\lambda^\chi = \lambda^\chi_x = \lambda^\chi_y$, and $\lambda^\Delta = \lambda^\Delta_x = -\lambda^\Delta_y$. We have also $n = n_i = \sum_\sigma \langle c^\dagger_{i\sigma} c_{i\sigma}\rangle$, $\chi_\tau = \chi_{ij}$, $\tau = x(y)$ for the bonds between the nearest-neighboring sites located along the x (y) direction, respectively, and $\Delta_\tau = \sqrt{2}\Delta_{ij}$. χ_{ij} and Δ_{ij} are given by Eq. 7.153, and all the above quantities are taken as real. Diagonalization of $\hat{H}_{R\lambda}$ (7.156), (in the present case denoted $\hat{H}^{(dSC)}_{R\lambda}$) yields

$$\hat{H}^{(dSC)}_{R\lambda}(\vec{A}, \vec{\lambda}) - \mu\hat{N} = \sum_{\mathbf{k}} E_{\mathbf{k}}(\hat{\gamma}^\dagger_{\mathbf{k}0}\hat{\gamma}_{\mathbf{k}0} + \hat{\gamma}^\dagger_{\mathbf{k}1}\hat{\gamma}_{\mathbf{k}1}) + \sum_{\mathbf{k}}(\xi_{\mathbf{k}} - E_{\mathbf{k}}) + C(\vec{A}, \vec{\lambda}), \tag{8.182}$$

with

$$E_{\mathbf{k}} = \sqrt{\xi^2_{\mathbf{k}} + D^2_{\mathbf{k}}}. \tag{8.183}$$

$\xi_{\mathbf{k}}$ and $D_{\mathbf{k}}$ appearing in (8.183) are given by

$$D_{\mathbf{k}} = \sqrt{2}\sum_\tau D_\tau \cos(k_\tau), \qquad \xi_{\mathbf{k}} = -2\sum_\tau T_\tau \cos(k_\tau) - \mu - \lambda. \tag{8.184}$$

Also, we have

$$T_\tau = -t_{1\tau} g^t_{1\tau} + \frac{3}{4} J_\tau g^J_\tau \chi_\tau + \lambda^\chi_\tau, \qquad D_\tau = \frac{3}{4} J_\tau g^J_\tau \Delta_\tau + \lambda^\Delta_\tau, \tag{8.185}$$

$$C(\vec{A}, \vec{\lambda}) = \Lambda\lambda n + \Lambda \sum_\tau \left(\frac{3}{4} J_\tau g^J_\tau (2\chi^2_\tau + \Delta^2_\tau) + 4\chi_\tau \lambda^\chi_\tau + 2\Delta_\tau \lambda^\Delta_\tau \right). \tag{8.186}$$

Below we use mainly RS (I) and (II). However, some results for RS IIIa and IIIb are also given for comparison, and in order to show notable differences between two versions of the formalism of Ref. [66]. The Gutzwiller renormalization factors read, respectively

$$g^t_{(I)\tau}(n, \chi_\tau, \Delta_\tau) = \frac{2(1-n)}{2-n}\left(1 - \frac{4\chi^2_\tau + 2\Delta^2_\tau}{(2-n)^2}\right), \qquad g^J_{(I)\tau}(n, \chi_\tau, \Delta_\tau) = \frac{4}{(2-n)^2},$$

$$g^{\Delta}_{(I)\tau}(n,\chi_\tau,\Delta_\tau)=\frac{2(1-n)}{2-n}\left(1+\frac{4\chi_\tau^2+2\Delta_\tau^2}{(2-n)^2}\right), \tag{8.187}$$

$$g^{t}_{(II)\tau}(n,\chi_\tau,\Delta_\tau)=\frac{2n(1-n)}{n(2-n)+4\chi_\tau^2},$$

$$g^{J}_{(II)\tau}(n,\chi_\tau,\Delta_\tau)=\frac{4n^2}{n^2(2-n)^2+(1-n)^2(4\Delta_\tau^2-8\chi_\tau^2)+4\Delta_\tau^4+16\chi_\tau^4}. \tag{8.188}$$

Note, that for RS (II), $g^{\Delta}_{\tau}=g^{t}_{\tau}$. Generalized Landau potential (7.158) reads

$$\mathcal{F}^{(dSC)}(\vec{A},\vec{\lambda})=C(\vec{A},\vec{\lambda})+\sum_{\mathbf{k}}\left((\xi_{\mathbf{k}}-E_{\mathbf{k}})-\frac{2}{\beta}\ln\left(1+e^{-\beta E_{\mathbf{k}}}\right)\right). \tag{8.189}$$

Using $\mathcal{F}^{(dSC)}(\vec{A},\vec{\lambda})$, we may easily obtain Eqs. (4.34), which now read

$$\begin{aligned}&\frac{\partial\mathcal{F}^{(dSC)}}{\partial n}=0,\quad \frac{\partial\mathcal{F}^{(dSC)}}{\partial\chi}=0,\quad \frac{\partial\mathcal{F}^{(dSC)}}{\partial\Delta}=0,\\ &\frac{\partial\mathcal{F}^{(dSC)}}{\partial\lambda}=0,\quad \frac{\partial\mathcal{F}^{(dSC)}}{\partial\lambda^{\chi}}=0,\quad \frac{\partial\mathcal{F}^{(dSC)}}{\partial\lambda^{\Delta}}=0.\end{aligned} \tag{8.190}$$

Solution of Eqs. (8.190) yields $\vec{A}_0=(n,\chi_0,\Delta_0)$ and $\vec{\lambda}_0=(\lambda_0,\lambda_0^{\chi},\lambda_0^{\Delta})$. Note, that the above equations are solved for μ, not for n, which is *a priori* known.

Similarly, within the (*s-c*) approach, the BdG equations (4.38) read now

$$\left(\frac{\partial\mathcal{F}^{(dSC)}}{\partial\lambda}\right)_{\vec{\lambda}=\vec{0}}=0,\quad \left(\frac{\partial\mathcal{F}^{(dSC)}}{\partial\lambda^{\chi}}\right)_{\vec{\lambda}=\vec{0}}=0,\quad \left(\frac{\partial\mathcal{F}^{(dSC)}}{\partial\lambda^{\Delta}}\right)_{\vec{\lambda}=\vec{0}}=0 \tag{8.191}$$

The optimal solution of (8.191) will be denoted $(n,\chi_{sc},\Delta_{sc})$.

8.1.1 Numerical results

We begin with the analysis of dSC solution at the 'magic doping', $x=0.125$. Equations (8.190) and (8.191) are solved numerically for the lattice of $\Lambda=\Lambda_x\Lambda_y$ sites, $\Lambda_x=\Lambda_y=256$, with $J_\tau=1$ $t_\tau=-3J$ and for low temperature, $\beta J=500$. Both dSC solution, as well as the isotropic normal state (N) (not discussed explicitly here) are present. For that particular value of doping the staggered flux (SF, cf. Subsection 8.2) state has not been found (*var*) or has been found unstable against N state (*s-c*).

In Tables I and II we provide values of thermodynamic potentials, chemical potential, and the optimal (equilibrium) values of mean-fields and Lagrange multipliers for dSC solution obtained within both methods (*var*, *s-c*) and different renormalization schemes (I, II, IIIa and IIIb).

Table I. Values of the thermodynamic potentials (per site) for dSC solutions. $\tilde{\Omega}$ (F) stands for $\Omega-\lambda N$ ($\Omega+\mu N$) for *var* and Ω_{s-c} ($\Omega_{s-c}+\mu_{sc}N$) for *s-c* methods, respectively.

Φ	*var* (I)	*var* (II)	*s-c* (I)	*s-c* (II)	(IIIa)	(IIIb)
Ω/Λ	-6.0444	-5.7587	-	-	-6.7953	-6.5457
$\tilde{\Omega}/\Lambda$	-1.0897	-1.0766	-1.0112	-1.0361	-1.1236	-1.1153
F/Λ	-1.3432	-1.3661	-1.3399	-1.3647	-1.2864	-1.3119

For fixed n, free energy F is the relevant thermodynamic potential. From the way the solutions are constructed, we may expect that the value of F obtained for a given RS is always lower within the *var* method than within the *s-c* one. This is fully confirmed by the obtained

numerical results. Also, value of F determines which solution of (8.190) or (8.191) corresponds to the stable equilibrium situation. However, free energy should *not* be used to favor one of the renormalization schemes. For example, if we compare the values of the F for *var* method in the present case, and also invoke its value F^0_{var}, corresponding to the simplest Gutzwiller factors (Eq. (7.170), cf. also Ref. [75]), we have that $F^0_{var} = -1.5070 < F^{II}_{var} < F^{I}_{var}$.[49] Obviously, it does not mean that RS defined by (7.170) should be preferred over RS (I) or RS (II). For the RMFT Hamiltonians the condition (4.123) is not fulfilled, i.e.,

$$\mathrm{Tr}\big[\hat{\rho}_\lambda(\vec{A}_0,\vec{\lambda}_0)\big\{\hat{H}_{tJ} - \hat{H}(\vec{A}_0,\vec{\lambda}_0)^{(dSC)}_{R\lambda}\big\}\big] \neq 0. \tag{8.192}$$

Consequently, the MF free-energy of the dSC solution cannot be regarded as an upper bound for the exact free energy of $\hat{H}_{tJ}$. Indeed, value F^0_{var} is much lower then the Variational Monte Carlo (VMC) result ($E_{VMC} \approx -1.33$) [75], hence it is very likely that F^0_{var} is also lower then the exact value (the latter remark is valid for other renormalization schemes as well). Also, if there was a good agreement between F_{var} obtained within the RMFT for the particular choice of RS, with the value of F obtained within the VMC approach, it would probably be accidental, and thus insignificant.[50] Values of thermodynamic potentials for RS of Ref. [66] are also given for comparison ((IIIa) and (IIIb)).

Table II. Values of the equilibrium chemical potentials and MF parameters ($\vec{A}_0$, $\vec{\lambda}_0$) for dRVB solutions. $\tilde{\mu}$ stands for $\lambda + \mu$ (*var*), and for μ_{sc} (*s-c*).

φ	*var* (I)	*var* (II)	*s-c* (I)	*s-c* (II)	(IIIa)	(IIIb)
μ	5.3729	5.0200	-	-	6.2959	5.9814
λ_0	-5.6625	-5.3509	-	-	-6.4820	-6.2061
$\tilde{\mu}_0$	-0.2897	-0.3309	-0.3757	-0.3755	-0.1860	-0.2247
χ_0	0.1941	0.1881	0.1901	0.1907	0.2001	0.1983
λ^χ_0	-0.1588	-0.1699	-	-	-0.2622	-0.2242
$\frac{1}{\sqrt{2}}\Delta_0$	0.1090	0.1320	0.1257	0.1235	0.0504	0.0788
$\frac{1}{\sqrt{2}}\lambda^\Delta_0$	-0.0892	-0.0111	-	-	-0.0706	-0.0984
$T_{\tau 0}$	0.8636	0.8487	1.0079	1.0228	0.7280	0.7837
$\sqrt{2}D_{x0}$	0.3383	0.5893	0.5957	0.5721	0.0598	0.1450

Let us now analyze Table II. First, we see that within RS (II), $|\lambda^\chi_{\tau 0}|$ is considerably larger than $|\lambda^\Delta_{\tau 0}|$, $\tau = x, y$. This indicates that the Hamiltonian (8.182) deviates from the Hartree-Fock form more with respect to mean-field variable χ_τ then with respect to Δ_τ.[51] This is due to the particular functional form of χ_τ-dependent, Δ_τ-independent renormalization of the hopping term. Also, the influence of this term is significant, since $|t| = 3J$. Similar conclusions are valid for RS (I), but then the χ_τ- and Δ_τ dependences of $\langle\hat{H}\rangle$ are more symmetric (cf. Eqs. (7.174) and (7.175), or (8.187), and thus the difference between $|\lambda^\chi_{\tau 0}|$ and $|\lambda^\Delta_{\tau 0}|$ is smaller then for RS (II).

Large value of $|\lambda^\chi_{\tau 0}|$ significantly affects (via T_τ and D_τ, Eq. (8.185)) the quasi-particle spectra, cf. Fig 2. Namely, the excitation energies, $E_{\mathbf{k}}$ (8.183), obtained within our method

[49] In Ref. [75], nonzero temperature ($\beta J = 500$) is also used.

[50] Let us note once more, that for such low value of temperature we use here, F is practically equal to the ground-state energy. This is why we compare the VMC results obtained for $T = 0$, with the RMFT results obtained for $T > 0$. In the $T \to 0$ limit, variational principle based on the Bogoliubov-Feynman inequality reduces to the variational principle of quantum mechanics. Therefore, the above discussion may be repeated for $T = 0$ without any changes, with the average in (8.192) being now evaluated with respect to the ground state of the MF Hamiltonian.

[51] We have divided Δ_τ by $1/\sqrt{2}$ in order to compare our results directly with the results of Refs. [73, 75], where slightly different definition of Δ_τ has been used.

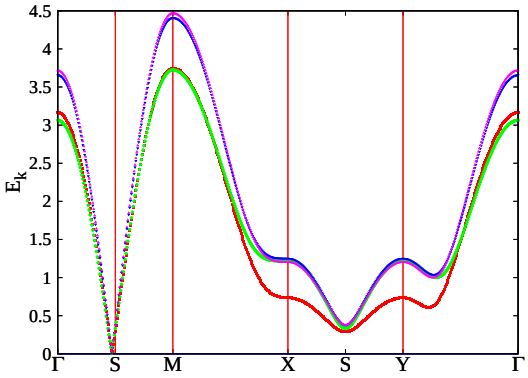

Figure 2: Dispersion relations along the main symmetry lines in the Brillouin zone for the dRVB solutions for a square lattice, of the size $\Lambda_x = \Lambda_y = 256$, and for the filling $n = 0.875$. Triangles - self-consistent, non-variational results for (I) and (II), red squares (I) and green circles (II)- the present variational method.

are always *lower* then those of the BdG self-consistent (*s-c*) approach (note, that $\lambda^{\chi}_{\tau 0} \cdot \chi_{\tau 0} < 0$ and $\lambda^{\Delta}_{\tau 0} \cdot \Delta_{\tau 0} < 0$). A related discussion is given in Ref. [66], point (2) of Summary (Section 7 of that Reference).

Although for RS (II) the differences between the methods (*var*, *s-c*) are pronounced mainly in the regions of the Brillouin zone which are not in the vicinity of the Fermi surface (for the superconducting state, the latter is defined by the condition $\xi_{\mathbf{k}} = 0$), the difference of the tangent at the cone near the point $S = (\frac{\pi}{2}, \frac{\pi}{2})$ may be of some significance. However, what is more important, RS (I) *var* yields the excitation energies which differ from those obtained in the remaining cases also along the X-Y direction, i.e., close to the Fermi surface. The reason for such behavior should be clear from the analysis of Table II. Along the X-Y direction $\xi_{\mathbf{k}} = \text{const}$ and the main contribution to $E_{\mathbf{k}}$ comes from $D_{\mathbf{k}}$, the value of which is determined in turn by D_x and D_y. The absolute values of the latter quantities are low for (I) *var*.

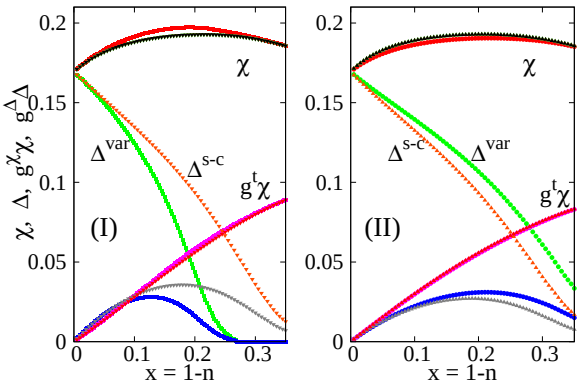

Figure 3: Renormalization schemes (I) and (II): Doping dependence of the optimal values of bond-order parameters $\chi_x = \chi_y$, the superconducting order parameters $\Delta_x = -\Delta_y$, as well as of their renormalized correspondents $g^t\chi_x = g^t\chi_y$ and $g^{\Delta}\Delta_x = -g^{\Delta}\Delta_y$, both for the *s-c* (triangles) and the *var* (squares) methods.

Doping dependence of mean-field quantities. Now we are going to discuss the changes appearing as a function of doping x. The x-dependences of the equilibrium values of mean-fields,

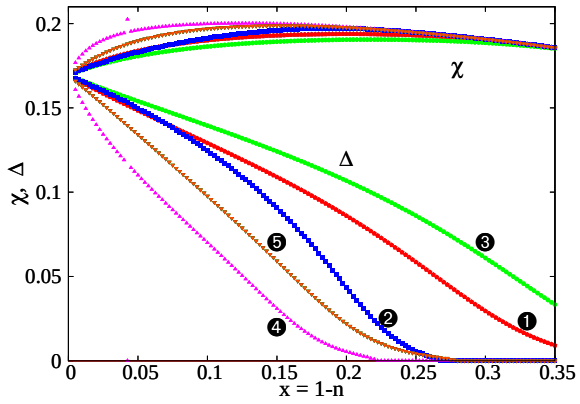

Figure 4: Doping dependence of the equilibrium values of the bond-order parameter (hopping correlation function) $\chi_x = \chi_y$ and the superconducting order parameter $\Delta_x = -\Delta_y$, for different renormalization schemes, and obtained within the *var* method. Data points: red (1)- the simplest renormalization scheme (Eqs. (7.170)), dark blue (2)- RS I (Eqs. (7.174)) and (7.175)), green (3)- RS II (Eqs. (7.176) and (7.177)), (4) and (5)- RS IIIa and RS IIIb, respectively (Eqs. (7.178) and (7.179)). Note the differences between IIIa and IIIb.

i.e., χ_0 and Δ_0, as well as the physical (renormalized) gap parameter $g^{\Delta}\Delta_0$ and renormalized hopping $g^t\chi$ for RS (I) and (II) and two methods of approach (*var*, *s-c*) are analyzed in Fig. 3. First, let us note, that $\chi_0(x)$ and $g^t(x, \chi_0(x), \Delta_0(x)) \cdot \chi_0(x)$ are quite similar for both renormalization schemes and methods. On the other hand, $g^{\Delta}(x, \chi_0(x), \Delta_0(x)) \cdot \Delta_0(x)$ and $\Delta_0(x)$ exhibit larger differences. In particular, differences between (I) *var* and (I) *s-c* are notable. Note also, that for (I) *var*, Δ_0 vanishes at the critical concentration $x_c \approx 0.27$. This is in a satisfactory agreement with the experimental results. Next, we analyze $\chi_0(x)$ and $\Delta_0(x)$

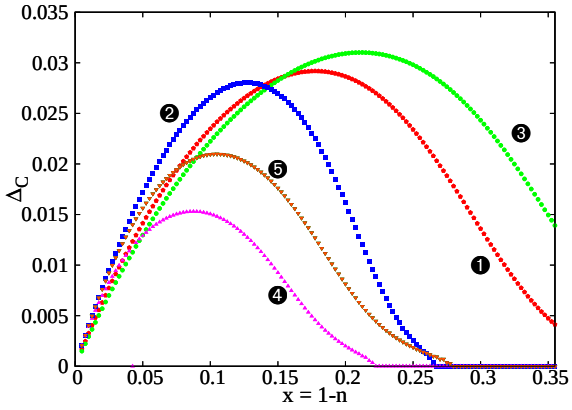

Figure 5: Doping dependence of the optimal value of the renormalized superconducting order parameter $\Delta_C = g^{\Delta}\Delta_x$, for different renormalization schemes, and obtained within the *var* method. Data points are labeled as in Fig. 4. Note the differences between IIIa and IIIb, in particular different value of the upper critical concentration x_c.

(Fig. 4) and $g^{\Delta}(x, \chi_0, \Delta_0) \cdot \Delta_0(x)$ (Fig. 5) within *var* method, for different renormalization schemes. For RS (I), (IIIa) and (IIIb), we obtain $0.21 < x_c < 0.27$, in the other cases we have $x_c > 0.35$.

Doping dependence of free energy and chemical potential. Doping dependence of the free energy, chemical potential and related mean-field quantities, λ and $\tilde{\mu} = \mu + \lambda$, is shown

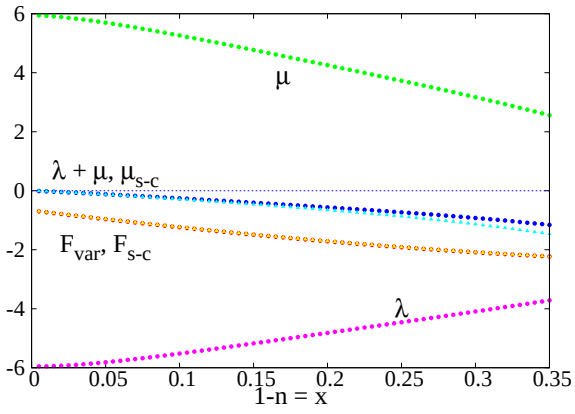

Figure 6: RS II: Doping dependences of the free energy (F_{var}, F_{s-c}), the chemical potentials μ, μ_{sc} as well as that of λ and $\tilde{\mu} = \lambda + \mu$ for the dSC state, obtained both within the present (var) and the (s-c) methods.

in Fig. 6, but only for RS (II). We emphasize once again, the chemical potential μ is the first derivative of F/Λ with respect to n, unlike in some of the previous mean-fields treatments [63, 68, 73, 75]. Note that neither $\tilde{\mu}$ nor λ have this property, as discussed in Subsection 4.7. Therefore, application of the standard thermodynamic identities and methods (e.g. the Maxwell construction) is probably not legitimate within such an approach, although this fact is sometimes ignored (cf. [78]).

8.2 Staggered flux solution

As a next example we analyze the staggered-flux (SF) state, first proposed by Affleck and Marston [114] as a variational trial MF state for the Heisenberg model on the square lattice. Later, the SF state has been investigated in a number of papers (cf. e.g. [62, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 194]), also because this state has been proposed as a candidate for the pseudogap phase in the cuprates [118]. SF state is also known to be one of most competitive (i.e., of lowest energy) states among the non-superconducting, non-magnetic ground states of the MF Hamiltonian (7.152) for a wide range of the values of its parameters [75]. Moreover, a state in which superconducting and SF orders coexist, has even lower energy [51]. Also, for $n = 1$, as a consequence of the local $SU(2)$ symmetry, SF state is equivalent to the RVB state of $d_{x^2-y^2}$ symmetry (as well as to many other MF states, e.g. $s + id$ wave state), [8, 58]. Therefore, even if SF state is not a superconducting state, it is very interesting in the context of the physics of cuprates, and deserves at least a brief analysis. However, here we are going to use SF state mainly as yet another convenient example, on which the differences between variational (var) and non-variational (s-c) solutions can be shown. We concentrate on RS (I) and (II), which, in the absence of superconducting order have quite similar form, and therefore the modifications caused by a method of solution (var, s-c) become dominant.

The SF state differs from the simplest Fermi sea (N) state by a presence of the complex hopping amplitude between the nearest neighboring sites

$$\chi_{ij} = |\chi| \exp((-1)^{(\mathbf{i_x}+\mathbf{j_y})} i\varphi) \equiv \xi_1 \pm i\xi_2. \tag{8.193}$$

Complex χ_{ij} imply the existence of circulating currents, direction of which changes from one elementary plaquette[52] to another in an alternating fashion. Consequently, a two-sublattice

[52] By an elementary plaquette we understand a square made of four sites (equivalently, bonds), which side is equal one lattice constant.

structure ('orbital antiferromagnet' [62]) emerges, with the unit cell of the size $\sqrt{2}a \times \sqrt{2}a$ (a is a lattice constant) in direct space and new folded Brillouin zone (NBZ).

Solving this problem within the framework of our method, we have to add appropriate constraints to the MF Hamiltonian according to Eq (7.156). This step introduces, apart from the molecular field λ coupled to n, also a complex Lagrange multiplier $\eta_{ij} = \eta_1 \mp i\eta_2$, connected to χ_{ij} (our sign convention for imaginary part of η_{ij} for each bond is opposite to that for χ_{ij}). Thus, we have three independent real mean fields $\vec{A} = (n, \xi_1, \xi_2)$, and three corresponding real Lagrange multipliers, $\vec{\lambda} = (\lambda, \eta_1, \eta_2)$. In the present case the renormalization factors (7.175) and (7.177) read, respectively

$$g^t_{(I)}(n,\xi_1,\xi_2) = \frac{2(1-n)}{2-n}\left(1 - \frac{4\,(\xi_1^2+\xi_2^2)}{(2-n)^2}\right), \qquad g^J_{(I)}(n,\xi_1,\xi_2) = \frac{4}{(2-n)^2}. \tag{8.194}$$

$$g^t_{(II)}(n,\xi_1,\xi_2) = \frac{2n(1-n)}{n(2-n)+4(\xi_1^2+\xi_2^2)},$$

$$g^J_{(II)}(n,\xi_1,\xi_2) = \frac{4n^2}{n^2(2-n)^2 - 8(1-n)^2(\xi_1^2+\xi_2^2) + 16(\xi_1^2+\xi_2^2)^2}. \tag{8.195}$$

Diagonalization of $\hat{H}_{R\lambda}$ (7.156) (in the present case denoted as $\hat{H}^{(SF)}_{R\lambda}$), yields

$$\hat{H}^{(SF)}_{R\lambda}(\vec{A},\vec{\lambda}) - \mu\hat{N} = C(\vec{A},\vec{\lambda}) + \sum_{\mathbf{k}s\sigma}^{NBZ} E_{\mathbf{k}s\sigma}\hat{\alpha}^\dagger_{\mathbf{k}s\sigma}\hat{\alpha}_{\mathbf{k}s\sigma}, \tag{8.196}$$

where $E_{\mathbf{k}s\sigma} = -\tilde{\mu} + s\sqrt{\epsilon^2_{\mathbf{k}} + \chi^2_{\mathbf{k}}}$, $\epsilon_{\mathbf{k}} = T_1\Gamma_+(\mathbf{k})$, $\chi_{\mathbf{k}} = T_2\Gamma_-(\mathbf{k})$, $\tilde{\mu} \equiv \mu + \lambda$, and for each $\mathbf{k}$, $s = \pm 1$ labels two sub-bands (two eigenvalues of the 2×2 matrix resulting from the Fourier transform. Operator $\hat{\alpha}^\dagger_{\mathbf{k}s\sigma}$ ($\hat{\alpha}_{\mathbf{k}s\sigma}$) creates (annihilates) quasiparticle with the quasimomentum $\mathbf{k}$, spin σ and the band index s. Also, $\Gamma_\pm(\mathbf{k}) \equiv 2(\cos(k_x) \pm \cos(k_y))$, and

$$T_1 = tg^t + \frac{3}{4}Jg^J\xi_1 + \eta_1, \qquad T_2 = \frac{3}{4}Jg^J\xi_2 + \eta_2, \tag{8.197}$$

$$C(\vec{A},\vec{\lambda}) = \Lambda\left(\lambda n + 3g^J(\xi_1^2+\xi_2^2) + 8(\xi_1\eta_1 + \xi_2\eta_2)\right). \tag{8.198}$$

Generalized Landau potential (7.158) takes now the following form

$$\mathcal{F}^{(SF)}(\vec{A},\vec{\lambda}) \;=\; C(\vec{A},\vec{\lambda}) - \frac{1}{\beta}\sum_{\mathbf{k}s\sigma}^{NBZ} \ln\left(1 + e^{-\beta E_{\mathbf{k}s\sigma}}\right). \tag{8.199}$$

Using $\mathcal{F}^{(SF)}(\vec{A},\vec{\lambda})$, we obtain Eqs. (4.34), which now read

$$\begin{aligned} &\frac{\partial\mathcal{F}^{(SF)}}{\partial n} = 0, \quad \frac{\partial\mathcal{F}^{(SF)}}{\partial\xi_1} = 0, \quad \frac{\partial\mathcal{F}^{(SF)}}{\partial\xi_2} = 0, \\ &\frac{\partial\mathcal{F}^{(SF)}}{\partial\lambda} = 0, \quad \frac{\partial\mathcal{F}^{(SF)}}{\partial\eta_1} = 0, \quad \frac{\partial\mathcal{F}^{(SF)}}{\partial\eta_2} = 0. \end{aligned} \tag{8.200}$$

Explicit form of the above equations will not be provided here. Solving Eqs. (8.200), we obtain $\vec{A}_0 \equiv (n, \xi_1^{(0)}, \xi_2^{(0)})$ and $\vec{\lambda}_0 = (\lambda_0, \eta_1^{(0)}, \eta_2^{(0)})$. As usual, for a given n the value of the chemical potential μ is obtained from Eqs. (8.200).

Similarity, within the (s-c) approach, we have to solve the following BdG equations (4.38)

$$\left(\frac{\partial\mathcal{F}^{(SF)}}{\partial\lambda}\right)_{\vec{\lambda}=\vec{0}} = 0, \quad \left(\frac{\partial\mathcal{F}^{(SF)}}{\partial\eta_1}\right)_{\vec{\lambda}=\vec{0}} = 0, \quad \left(\frac{\partial\mathcal{F}^{(SF)}}{\partial\eta_2}\right)_{\vec{\lambda}=\vec{0}} = 0. \tag{8.201}$$

Solution of (8.201) will be denoted $\vec{A}_{sc} \equiv (n, \xi_1^{(sc)}, \xi_2^{(sc)})$.

8.2.1 Numerical results

As mentioned in Section 8.1, at $x = 1/8$ the existence of SF solution ($\xi_2^{(0)} \neq 0$) of Eqs. (8.200) has not been numerically confirmed for the *var* method, whereas for the *s-c* method the SF solutions of Eq. (8.201) has higher energy than the Fermi sea (N) solution with $\xi_2^{(0)} = 0$. We have found that in all four cases the SF $\rightarrow$ N transition is located at the critical concentration $x_c < 0.12$. It should be noted here, that by using the simplest renormalization scheme given by (7.170), one obtains $x_c \sim 0.15$ (cf. Fig. 2 b of Ref. [75]). However, our numerical procedures for the *var* method turned out to be unstable in the vicinity of the x_c. For that reason, we chose doping $x = 13/128$ ($n \approx 0.898$), which is in a safe distance from x_c obtained within each method, but for which the differences between the methods (generally increasing with the increasing doping) are still pronounced. The parameters of the Hamiltonian are the same as for the RVB case, except sign convention for t, now $t = 3$ ($t = -3$ in the RVB case). We also use larger lattice, $\Lambda_x = \Lambda_y = 512$. Again, we work with low $k_B T/J = 1/(\beta J) = 1/500$.

Within the *s-c* approach the (fictitious) flux is defined as $\Phi_\Box = \frac{1}{2\pi} \sum_{\langle ij \rangle \in \Box} \mathrm{Arg}(\chi_{ij})$, where $\Box$ denotes elementary plaquette (cf. e.g. [75]). Also, for the *s-c* method with $\eta_1 = \eta_2 = 0$, we have

$$\mathrm{Arg}(\chi_{ij}) = \arctan\left(\frac{\xi_2}{\xi_1}\right) = \arctan\left(\frac{T_2}{T_1 - tg^t}\right), \tag{8.202}$$

the last equality follows from Eqs. (8.197). Interestingly, this equality holds also within the variational approach with $\eta_1 \neq 0$, $\eta_2 \neq 0$. Namely, it can be shown, that $\xi_1^{(0)}/\xi_2^{(0)} = \eta_1^{(0)}/\eta_2^{(0)}$, which together with (8.197) yields (8.202). Thus the two possible and *a priori* different definitions of $\Phi_\Box$ within *var* method turn out to be equivalent. Antiferromagnetic correlations are defined on the MF level as

$$S_{AF} = -\frac{3}{2} g^J (\xi_1^2 + \xi_2^2), \tag{8.203}$$

cf. Refs. [68, 75]. Thermodynamic potentials, the optimal values of the mean-field variables and related quantities for the SF solution are listed in Tabs. III and IV.

Table III. Equilibrium values of the thermodynamic potentials (per site) for SF solutions, for $n \approx 0.898$. $\tilde{\Omega}$ (F) stands for $\Omega - \lambda N$ ($\Omega + \mu N$) for *var* and Ω_{s-c} ($\Omega_{s-c} + \mu_{sc} N$) for *s-c* methods, respectively.

Φ	*var* (I)	*var* (II)	*s-c* (I)	*s-c* (II)
Ω/Λ	-5.90103948	-5.75555848	-	-
$\tilde{\Omega}/\Lambda$	-0.69003640	-0.71583672	-0.49006797	-0.49344983
F/Λ	-1.18762800	-1.16898668	-1.18536044	-1.16599431

Table IV. Values of chemical potentials, optimal MF parameters, and related quantities for SF solutions at $n \approx 0.898$. $\tilde{\mu}$ stands for $\lambda + \mu$ (*var*), and for μ_{sc} (*s-c*).

φ	*var* (I)	*var* (II)	*s-c* (I)	*s-c* (II)
μ	5.24623	5.10505	-	-
λ	-5.80007	-5.60943	-	-
$\tilde{\mu}$	-0.55384	-0.50438	-0.77389	-0.74857
ξ_1	0.19222	0.19321	0.18805	0.18844
ξ_2	0.10298	0.09840	0.11856	0.11731
η_1	-0.13476	-0.14848	-	-
η_2	-0.07220	-0.07562	-	-
S_{AF}	-0.23514	-0.22522	-0.24436	-0.23527
$\Phi_\Box$	0.31311	0.29987	0.35811	0.35449
T_1	0.80695	0.77919	0.92799	0.91123
T_2	0.18241	0.16008	0.29312	0.28009

From Tabs. III and IV we see, that in contrast to the dSC case, now the differences between renormalization schemes ((I) vs (II)) within each (*var*, *s-c*) method are small. This is because in the absence of Δ_{ij}, the χ_{ij} -dependences of $g^t_{(I)}$ and $g^t_{(II)}$ are quite similar, and the χ_{ij} -dependence of $g^J_{(II)}$ is weak for doping $x \approx 0.1$, thus causing no qualitative and only minor quantitative differences between the renormalization schemes. On the other hand, the generic modifications introduced by the variational approach within each RS are more significant.

Namely, from Tab. IV we see, that the *s-c* method favors SF solution as compared the *var* method. This is indicated by smaller values of $\Phi_\Box$ and ξ_2 obtained within the latter approach. The two latter quantities, or some combination of them, may serve as an order parameter for the SF $\leftrightarrow$ N transition. Also, T_1 and T_2 (8.197), which determine the quasi-particle spectra, are smaller within the *var* method, and so are the quasi-particle energies.

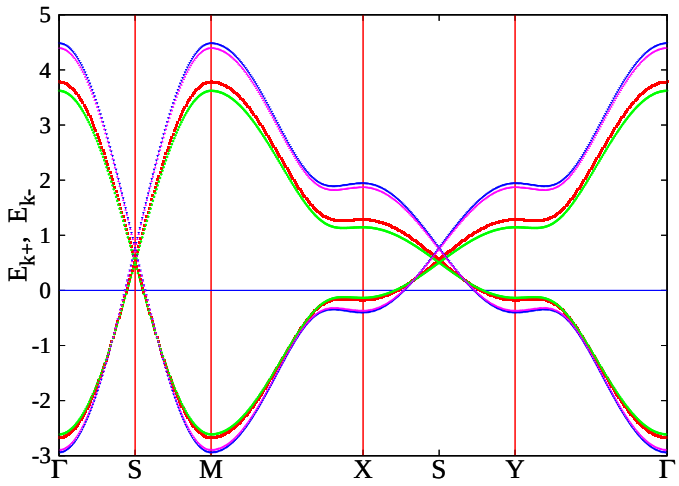

Figure 7: Dispersion relations for both upper ($E_{\mathbf{k}+}$) and lower ($E_{\mathbf{k}-}$) subbands for the SF solutions along the main symmetry lines in the Brillouin zone of the square lattice. The lattice is of the size $\Lambda = \Lambda_x \Lambda_y$, where $\Lambda_x = \Lambda_y = 512$, and the filling is $n \approx 0.898$. Triangles- the self-consistent, non-variational results for (I) and (II), red squares (I) and green circles (II)- the present variational method. Explicitly, for ($E_{\mathbf{k}+}$) near the maximum at point M, from the bottom to the top: (II) *var*, (I) *var*, (II) *s-c* (violet triangles), (I) *s-c* (dark blue triangles).

Next, let us analyze the dispertion relation (Fig. 7). The characteristic feature of the SF solution are *Dirac cones*,[53] i.e., regions of the FBZ, where the quasiparticle energy depends linearly on the absolute value of quasi-momentum, $E_{\mathbf{k}\sigma s} = v_{\sigma s}(\hat{\mathbf{k}})|\mathbf{k}|$ ($\hat{\mathbf{k}}$ is the unit vector). In contrast to the RVB case, now the Dirac cones are pinned exactly at S $= (\pi/2, \pi/2)$ (cf. Fig. 1 of Ref. [75]). Please note, that Fermi velocities in both the nodal ($(0,0) \rightarrow (\pi,\pi)$) and anti-nodal ($(0,\pi) \rightarrow (\pi,0)$) directions are smaller within the *var* method.

8.3 Pomeranchuk instability

So far we have focused on the solutions for which the x and y directions are equivalent.[54] However, a spontaneous (i.e., occurring for the x-y symmetric choice of the parameters in the microscopic Hamiltonian, $t_x = t_y$, $J_x = J_y$, and not accompanied with the lattice distortion) breakdown of the equivalence of the x- and y- directed correlations is possible already in the

[53] This term follows from a formal analogy of such spectrum with that of the massless Dirac fermions.

[54] As mentioned in the Introduction, the $d_{x^2-y^2}$ symmetry of the superconducting gap, which imply that $\Delta_x = -\Delta_y$ does not violate x-y symmetry, because $\mathcal{F}^{(dSC)}$ (8.189) is an even function of both Δ_x and Δ_y.

normal phase. This is termed the *Pomeranchuk instability* (PI), [125, 126, 127, 128, 129, 130] that manifests itself by lowering of the discrete C_{4v} symmetry of the Fermi surface.

It is important to note at this point, that the *isotropic* ($|\Delta_x| = |\Delta_y|$) superconducting solution (dSC) analyzed in Subsection 8.1 is energetically favorable as compared the anisotropic normal solution analyzed here. The latter is obtained only if the superconducting order is suppressed 'by hand' [33]. Also, for the square lattice and x-y symmetric Hamiltonian, the solution with $|\Delta_x| \neq |\Delta_y|$ has not been found.

Even if PI may seem to be auxiliary from the point of view of the present Thesis (no pairing), we would like to analyze it briefly, for the following reasons. First, PI is another good example of the nontrivial differences between the two methods analyzed here (*var*, *s-c*) of solving the MF models. Second, PI is a phenomenon that recently gained a considerable interest, and for which the t-J model is frequently used as a starting point of the analysis. Moreover, lack of a discrete rotational symmetry (in that case not spontaneous, but caused by the orthorhombic lattice distortion) may appear also in the superconducting state. This latter topic, as well as a more complete analysis of the PI in a normal state will be analyzed elsewhere [212]. Here we present only our earlier results [82] obtained using RS (II).

In the present case, $\vec{A} = (n, \chi_x, \chi_y)$, and $\vec{\lambda} = (\lambda, \lambda^\chi_x, \lambda^\chi_y)$. MF Hamiltonian and generalized Landau potential $\mathcal{F}^{(PI)}(\vec{A}, \vec{\lambda})$ we use here can be obtained from (8.182) and (8.189), by putting $\Delta_x = \Delta_y = 0$, and $\chi_x \neq \chi_y$, respectively. Consequently, within the *var* method we solve the following set of equations

$$\begin{aligned} &\frac{\partial \mathcal{F}^{(PI)}}{\partial n} = 0, \quad \frac{\partial \mathcal{F}^{(PI)}}{\partial \chi_x} = 0, \quad \frac{\partial \mathcal{F}^{(PI)}}{\partial \chi_y} = 0, \\ &\frac{\partial \mathcal{F}^{(PI)}}{\partial \lambda} = 0, \quad \frac{\partial \mathcal{F}^{(PI)}}{\partial \lambda^\chi_x} = 0, \quad \frac{\partial \mathcal{F}^{(PI)}}{\partial \lambda^\chi_y} = 0, \end{aligned} \tag{8.204}$$

whereas for the *s-c* case we have the corresponding BdG equations (4.38)

$$\left(\frac{\partial \mathcal{F}^{(PI)}}{\partial \lambda}\right)_{\vec{\lambda}=\vec{0}} = 0, \quad \left(\frac{\partial \mathcal{F}^{(PI)}}{\partial \lambda^\chi_x}\right)_{\vec{\lambda}=\vec{0}} = 0, \quad \left(\frac{\partial \mathcal{F}^{(PI)}}{\partial \lambda^\chi_y}\right)_{\vec{\lambda}=\vec{0}} = 0. \tag{8.205}$$

8.3.1 Numerical results

In Fig. 8 (a) the doping dependences of the bond-order parameters χ_x and χ_y are displayed for the x-y symmetric (N) and the symmetry-broken (PI) solution, both within the present (χ^{var}_τ) and the standard (χ^{s-c}_τ) methods. Within the *s-c* scheme, PI solution has been found up to $x \approx 0.091$. However, a comparison of the respective free-energy differences, $\Delta F_{s-c} \equiv F^N_{s-c} - F^{PI}_{s-c}$ and $\Delta F_{var} \equiv F^N_{var} - F^{PI}_{var}$, (cf. Fig. 8 (a)) reveals that this solution becomes unstable against the N state for $x \approx 0.021$. Therefore the PI $\rightarrow$ N phase transition is certainly discontinuous. On the other hand, it seems that within the *var* treatment, the PI solution does not exist for $x > x^{var}_c \approx 0.044$, where $\Delta F_{var} \approx 0$, in qualitative agreement with what is expected for the continuous phase transition.[55] Nonetheless, from the above analysis it is clear, that for PI those two methods of approach (*s-c*, *var*) yield qualitatively different predictions.

We have searched for a SF state with the spontaneously broken x-y symmetry, i.e., a coexistence of PI and standard SF phase. However, this solution has not been found.

Within our method, we have also examined the ladder system, analyzed in Ref. [63]. The ladder geometry explicitly breaks the x-y symmetry, and the superconducting ground state

[55] Within the numerical analysis, due to the unavoidable numerical errors in the close vicinity of the phase transition, we cannot unambiguously identify the order of the phase transition. This problem may be probably bypassed by plotting the Landau potential $\mathcal{F}_z$ (cf. Section 4.5). This would allow us to infer the order of the phase transition in a way simpler than a direct analysis of the x-dependence of the order parameter.

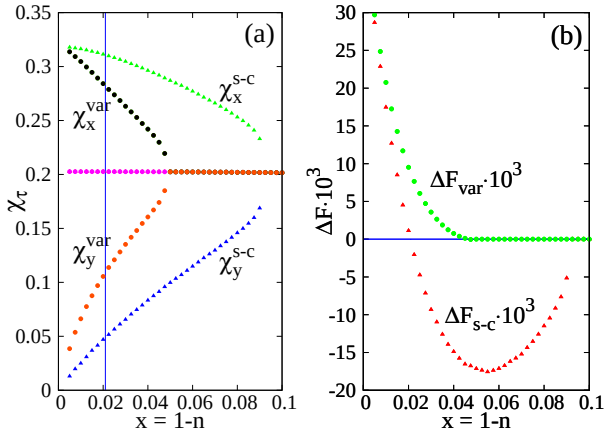

Figure 8: Renormalization scheme (II): doping dependence of bond-order parameters χ_x and χ_y (left), and the free-energy differences ΔF (right) both for x-y symmetric (N) and the x-y symmetry-broken states (Pomeranchuk instability, PI) within both the present (*var*, filled circles) and the non-variational (*s-c*, triangles) approach, respectively. The vertical line marks the phase transition within the *s-c* method. For details, see main text.

is found, for which $|\Delta_x| \neq |\Delta_y|$. We do not present the results here, not however, that the modifications introduced by the *var* approach are qualitatively similar to the those analyzed so far for the dSC and SF states. For example, the excitation energies $E_{\mathbf{k}}$ are lower within the *var* method, as compared to the *s-c* approach, analyzed in Ref. [63].

9 Results II: Optimal renormalization scheme and its application to t-J model

In the previous Section, we have compared results obtained by using various renormalization schemes (RS). Now, we apply the RS of Ref. [76] to study different forms of the t-J Hamiltonian (6.126). Also, the role of the next-nearest (t') and next-next-nearest (t'') neighbor hopping integrals is analyzed. The non-variational method based on the BdG equations (7.160) will no longer be utilized. Consequently, it is convenient to apply an alternative formulation of RMFT, as given by Eqs. (7.166) - (7.169) (Section 7.4.3). This Section provides the principal results of this Thesis for the t-J model.

9.1 Non-standard formulation of RMFT approach

9.1.1 RMFT Hamiltonian

Our point of departure is the RMFT Hamiltonian (7.164)

$$\begin{aligned}\hat{H}^{(\sim)}_{R\lambda} &= -\sum_{\langle ij\rangle\sigma}\Big(\tilde{\eta}_{ij\sigma}\big(c^{\dagger}_{i\sigma}c_{j\sigma}-\chi_{ij\sigma}\big)+\text{H.c.}\Big)-\sum_{i\sigma}\tilde{\lambda}^{(n)}_{i\sigma}\big(\hat{n}_{i\sigma}-n_{i\sigma}\big)\Big)\\ &- \sum_{\langle ij\rangle}\Big(\tilde{\gamma}_{ij}\big(\hat{\Delta}_{ij}-\Delta_{ij}\big)+\text{H.c.}\Big)+W(\chi_{ij\sigma},\Delta_{ij},n_{i\sigma}).\end{aligned}$$

In above, $\hat{\Delta}_{ij}$ is given by Eqs. (7.165). To obtain an explicit form of $W(\chi_{ij\sigma},\Delta_{ij},n_{i\sigma}) = \langle\hat{H}_{tJ}\rangle^{app}_{C}$ appearing in (7.163) and (7.164), we assume that the following mean fields and the corresponding Lagrange multipliers may take non-zero values: $\chi_{ij}=\xi_s$ and $\tilde{\eta}_{ij\sigma}=\tilde{\eta}_{ij}\equiv\eta_s$, with $s=1,3,5$ corresponding to sites with $d(i,j)\equiv|\vec{R}_i-\vec{R}_j|$ equal 1, $\sqrt{2}$, and 2 (in units of the lattice constant a), respectively; $\Delta_{ij}=\langle\hat{\Delta}_{ij}\rangle=\Delta_{x(y)}$, with $\Delta_{x(y)}=\pm\Delta$ and $\tilde{\gamma}_{ij}=\gamma_{x(y)}=\pm\gamma$ ($d_{x^2-y^2}$ symmetry), both nonzero for $d(i,j)=1$ and zero otherwise, as well as $n_{i\sigma}=n_{i\bar{\sigma}}=n/2$ and $\tilde{\lambda}^{(n)}_{i\sigma}=\lambda$. Therefore, we have

$$\vec{A}=(\xi_1,\xi_3,\xi_5,\Delta,n),\qquad \vec{\lambda}=(\eta_1,\eta_3,\eta_5,\gamma,\lambda). \tag{9.206}$$

Then, $W(\vec{A})$ is obtained by applying formalism of Ref. [76] (for more details, please consult Appendix D (Subsection 11.4) and the original paper). Explicitly, we have

$$W(\vec{A})=W_t+W_J+W_3, \tag{9.207}$$

where

$$\frac{W_t}{\Lambda}=\frac{16(1-n)}{2-n}\left\{t_1\xi_1\left(1-\frac{4(\Delta^2+\xi_1^2)}{(2-n)^2}\right)+\sum_{s=3,5}t_s y_s\xi_s\left(1-\frac{4\xi_s^2}{(2-n)^2}\right)\right\}, \tag{9.208}$$

$$\frac{W_J}{\Lambda}=-4J\left(\frac{3\left(\Delta^2+\xi_1^2\right)+c_1(1-n)^2\left(\Delta^2-\xi_1^2\right)}{(2-n)^2}\right)-Jc_1n^2/2, \tag{9.209}$$

and

$$\begin{aligned}
\frac{W_3}{\Lambda} &= -32c_2J\frac{(1-n)}{(2-n)^3}\cdot\Big\{\Big(1-\frac{n^2}{4}\Big)\xi_1^2+(2-n)\xi_1^2\xi_3-\frac{n\xi_3^3}{2}\\
&+ n\xi_3\Delta^2+\frac{1}{2}n\xi_3\Big(1-\frac{n}{2}\Big)^2-\Big(1-\frac{n}{2}\Big)\Big(2-\frac{n}{2}\Big)\Delta^2\Big\}\\
&- 16c_2J\frac{(1-n)}{(2-n)^3}\cdot\Big\{\Big(1-\frac{n^2}{4}\Big)\xi_1^2+(2-n)\xi_1^2\xi_5-\frac{n\xi_5^3}{2}\\
&+ n\xi_5\Delta^2+\frac{1}{2}n\xi_5\Big(1-\frac{n}{2}\Big)^2+\Big(1-\frac{n}{2}\Big)\Big(2-\frac{n}{2}\Big)\Delta^2\Big\}\\
&+ c_2\frac{t_3J(1-n)}{|t_1|\left(1-\frac{n}{2}\right)^3}\cdot\Big\{\xi_1\Big(8\Big(1-\frac{n}{2}\Big)\left(\xi_1^2+\xi_3^2\right)+4n\left(\Delta^2-\xi_1^2\right)\Big)\\
&+ 16\Big(\frac{n}{2}+1\Big)\Big(1-\frac{n}{2}\Big)\xi_1\xi_3+4n\Big(1-\frac{n}{2}\Big)^2\xi_1\Big\} \qquad (9.210)
\end{aligned}$$

In the above formulas, Λ is the number of lattice sites. Parameters y_3 and y_5 in W_t (9.208), equal 0 or 1, allow us to select the way in which the hopping integrals to the more distant neighbors are renormalized. All presented results are for $y_3 = y_5 = 1$, but we comment also on the situation with either y_3 or y_5 equal zero. The model parameters are J, $t_1 = t, t_3 = t'$ and $t_5 = t''$. Note, that in $\hat{H}_{tJ}$ (6.126) we retain all terms of order of t^2/U ($\sim J$) and tt'/U ($\sim J' \equiv Jt_3/|t_1|$), and neglect a smaller terms $\sim (t')^2/U$ ($\sim J'' \equiv J(t_3/t_1)^2$). As a consequence, we obtain the following MF grand Hamiltonian

$$\begin{aligned}
\hat{K}_\lambda^{(\alpha)} &\equiv -\sum_{\langle ij\rangle_s}\sum_\sigma\sum_s\Big(\tilde{\eta}_s\big(c_{i\sigma}^\dagger c_{j\sigma}-\xi_s\big)+\text{H.c.}\Big)\\
&- \sum_{\langle ij\rangle}\Big(\tilde{\gamma}\big(\hat{\Delta}_{ij}-\Delta_\tau\big)+\text{H.c.}\Big)-\mu\sum_{i\sigma}\hat{n}_{i\sigma}\\
&- \sum_{i\sigma}\big(\tilde{\lambda}_{i\sigma}^n\big(\hat{n}_{i\sigma}-n/2\big)\Big)+W(\xi_s,\Delta_\tau,n). \qquad (9.211)
\end{aligned}$$

Next, we diagonalize $\hat{K}_\lambda$ (7.164) via the Bogoliubov-Valatin transformation. This step yields

$$\hat{K}_\lambda^{(\alpha)} = \sum_{\mathbf{k}} E_{\mathbf{k}}(\hat{\gamma}_{\mathbf{k}0}^\dagger\hat{\gamma}_{\mathbf{k}0}+\hat{\gamma}_{\mathbf{k}1}^\dagger\hat{\gamma}_{\mathbf{k}1})+\sum_{\mathbf{k}}(\xi_{\mathbf{k}}-E_{\mathbf{k}})+C^{(\alpha)}(\vec{A},\vec{\lambda}), \qquad (9.212)$$

with

$$C^{(\alpha)}(\vec{A},\vec{\lambda}) = W(\vec{A},\vec{\lambda})+\Lambda(8\sum_s\xi_s\eta_s+4\Delta\gamma+\lambda n). \qquad (9.213)$$

Also, we have

$$E_{\mathbf{k}} = \sqrt{\xi_{\mathbf{k}}^2+D_{\mathbf{k}}^2}, \quad D_{\mathbf{k}} = -\frac{\gamma}{2}\Gamma_-(\mathbf{k}),$$

$$\xi_{\mathbf{k}} = -\left(\eta_1\Gamma_+(\mathbf{k})+\eta_3\Theta(\mathbf{k})+\eta_5\Gamma_5(\mathbf{k})+\tilde{\mu}\right), \qquad (9.214)$$

with

$$\Gamma_\pm(\mathbf{k}) = 2(\cos(k_x)\pm\cos(k_y)), \quad \Theta(\mathbf{k}) = 4\cos(k_x)\cos(k_y),$$

$$\Gamma_5(\mathbf{k}) = 2(\cos(2k_x)+\cos(2k_y)), \quad \text{and} \quad \tilde{\mu} = \mu+\lambda. \qquad (9.215)$$

9.1.2 Generalized Landau potential and equations (4.34)

Using $\hat{K}_\lambda^{(\alpha)}$, we obtain the grand-canonical density operator (7.169), i.e., $\hat{\rho}_\lambda = \mathcal{Z}_\lambda^{-1} \exp(-\beta \hat{K}_\lambda^{(\alpha)})$. Consequently, generalized Landau potential (4.33) has the form

$$\mathcal{F} = \mathcal{F}^{(\alpha)} \equiv C^{(\alpha)} + \sum_{\mathbf{k}} \big((\xi_{\mathbf{k}} - E_{\mathbf{k}}) - \frac{2}{\beta} \ln \big(1 + e^{-\beta E_{\mathbf{k}}} \big) \big). \tag{9.216}$$

Eqs. (4.34), i.e., the necessary conditions for $\mathcal{F}^{(\alpha)}$ to have a minimum subject to constraints read now

$$\partial_w \mathcal{F}^{(\alpha)} = \partial_w C^{(\alpha)} = \vec{0}_5, \tag{9.217}$$

$$\partial_z \mathcal{F}^{(\alpha)} = \vec{0}_5, \tag{9.218}$$

where $w \in \{\xi_1, \xi_3, \xi_5, \Delta, n\}$ and $z \in \{\eta_1, \eta_3, \eta_5, \gamma, \lambda\}$. In contrast to the formulation used in Section 8, equations (9.217) may be easily solved in an analytic fashion. However, solutions of the remaining Eqs. (9.218) must be determined numerically. The optimal values of mean fields and Lagrange multipliers will be denoted $\vec{A}_0 = (\xi_1^{(0)}, \xi_3^{(0)}, \xi_5^{(0)}, \Delta^{(0)}, n)$ and $\vec{\lambda}_0 = (\eta_1^{(0)}, \eta_3^{(0)}, \eta_5^{(0)}, \gamma^{(0)}, \lambda^{(0)})$, respectively.

Note, that within the present formulation, the quasiparticle energies $E_{\mathbf{k}}$ (9.214) depend functionally only on the Lagrange multipliers. Nonetheless, the latter become explicitly known functions of mean-fields after Eqs. (9.217) are solved.

9.1.3 Characteristics of the model: qualitative analysis

The analytical expression for $W(\vec{A})$ (9.207) allows us to make some qualitative predictions concerning the properties of our MF model before any numerical analysis is carried out. This is because at low T, the entropic part of the free energy is small, and therefore

$$F \approx U = W(\vec{A}_0) = \big\langle \hat{K}^{(\alpha)}(\vec{A}_0) \big\rangle + \mu N. \tag{9.219}$$

First, from (9.208) we expect a strong tendency to the superconductivity suppression for higher doping, as SC order leads to the band energy decrease $\sim \Delta^2$. On the other hand, in that regime the renormalized band energy becomes predominant over the exchange part. In effect, the normal state is favored over SC for $x > x_c$ with x_c smaller than obtained within previous MF treatments [33, 39, 58, 61]. Second, from (9.209) we infer, that the influence of the $\hat{\nu}_i \hat{\nu}_j/4$ term on W_J is small except for the largest doping; this is due to the presence of the $(1-n)^2$ pre-factor (the other term $\sim n^2$ merely shifts the chemical potential). On the other hand, W_3 (9.210) is multiplied only by $(1-n) = x$ pre-factor. Consequently, for higher x this term becomes rather important. This is also due to the large number of distinct three-site terms present for a given initial site and spin direction (eight for $d(i,j) = \sqrt{2}$, and four for $d(i,j) = 2$). Also, this part of W is expected to suppress the SC order, as the term $\sim \Delta^2$ in W_3 contains a factor, which is positive for reasonable values of other mean fields. Therefore, the energy gain of forming the superconducting state is *lower* if the $\hat{H}_3$ term is present.

9.2 Numerical results

After Eqs. (9.217) are solved analytically for Lagrange multipliers, we solve numerically the remaining Eqs. (9.218) for the mean fields. Again, we use periodic boundary conditions on the lattice of $\Lambda = \Lambda_x \Lambda_y = 512^2$ sites, to minimize finite size effects, and assume that $k_B T = 2 \cdot 10^{-3} J$, (practically equivalent to $T = 0$). Also, in the most cases we take the parameters $|t|/J = 3$ (corresponding to $U/|t| = 12$ for the Hubbard model), $t'/t = 0$ or -0.25, and $t'' = 0$ ($t'' \neq 0$ is

consider at the end of Sub-subsection 9.2.2, and $|t|/J \neq 3$ in Sub-subsection 9.2.3). Additionally, we take either $|t| = 0.3$ eV or $|t| = 0.4$ eV, which correspond roughly to the lower and the upper limits of the realistic values of this parameter, depending on the compound. Values of $|t|$ close to 0.4 eV have been determined from the band-structure calculations, [8, 213], whereas $|t| = 0.3$ eV is used e.g. in Refs. [61] and [48].

To highlight the influence of various forms of the t-J Hamiltonian (6.126), the results for different values of c_1 and c_2 are analyzed. The numbers 1, 2, 3 (4, 5, 6) in Figs. 9-12 correspond to the three separate situations: $c_1 = c_2 = 0$, $c_1 = 1$ and $c_2 = 0$, and $c_1 = c_2 = 1$, each taken for $t'/t = 0$ ($t'/t = -0.25$), respectively. Explicitly,

$$\begin{array}{lll} (1) & c_1 = 0,\, c_2 = 0, & t' = 0, \\ (2) & c_1 = 1,\, c_2 = 0, & t' = 0, \\ (3) & c_1 = 1,\, c_2 = 1, & t' = 0, \\ (4) & c_1 = 0,\, c_2 = 0, & t' = -t/4, \\ (5) & c_1 = 1,\, c_2 = 0, & t' = -t/4, \\ (6) & c_1 = 1,\, c_2 = 1, & t' = -t/4. \end{array}$$

9.2.1 Hole concentration $x = 0.175$

Optimal values of mean field parameters. In Table V we detail the equilibrium values of the mean fields and Lagrange multipliers for cases 1-6, for $x = 0.175$, the representative hole concentration in the overdoped regime. For this particular value of x, in all six cases we obtain superconducting solution with $\Delta \neq 0$.

Table V. Equilibrium values of the MF parameters, for $x = 0.175$.

φ	1	2	3	4	5	6
$\xi_1^{(0)}$	0.1970	0.1969	0.1990	0.1924	0.1922	0.1944
$\xi_3^{(0)}$	0.0468	0.0465	0.0505	0.0241	0.0239	0.0225
$\xi_5^{(0)}$	-0.0080	-0.0076	-0.0144	0.0337	0.0340	0.0383
$\Delta_x^{(0)}$	0.0687	0.0708	0.0202	0.0903	0.0919	0.0534
$\eta_1^{(0)}$	1.0080	1.0030	1.2355	1.0031	0.9982	1.1845
$\eta_3^{(0)}$	0.0000	0.0000	0.0803	-0.2223	-0.2223	-0.2118
$\eta_5^{(0)}$	0.0000	0.0000	0.0408	0.0000	0.0000	0.0404
$\gamma_x^{(0)}$	0.1584	0.1665	0.0320	0.2126	0.2205	0.0834
$\tilde{\mu}$	-0.4069	-0.4080	-0.2935	-0.8633	-0.8614	-0.9406

A closer look at Table V indicates that presence of the term $\sim c_1$ in the t-J Hamiltonian has practically no effect. However, the influence of both nonzero t_3, as well as the three-site terms $\sim c_2$ is substantial. Also, let us note that if ξ_3 (ξ_5) does not appear in $W(\vec{A})$ (9.207), we have $\eta_3^{(0)} = 0$ (cases 1 and 2) and $\eta_5^{(0)} = 0$ (cases 1, 2, 4, 5), which follows from Eqs. (9.217). However, in all cases 1-6 we still have $\xi_3^{(0)} \neq 0$ and $\xi_5^{(0)} \neq 0$. This is in agreement with the discussion provided in Section 7, see 7.4.4. Let us also note, that for cases 3 and 6, $W(\vec{A})$ depends in a non-trivial way on both ξ_3 and ξ_5, even if $t_3 = 0$ or $t_5 = 0$. This is because the three-site terms introduce an effective hopping to the next- and second-next nearest neighbor, even if such hopping is absent in the original Hubbard model.

Quasiparticle dispersion. For the parameters listed in Table V, in Fig. 9 we plot the dispersion relation of the Bogoliubov quasiparticles. At this particular doping ($x = 0.175$), the influence of $\hat{H}_3$ on $E_{\mathbf{k}}$ is of comparable magnitude to that of having nonzero t'. Also, although

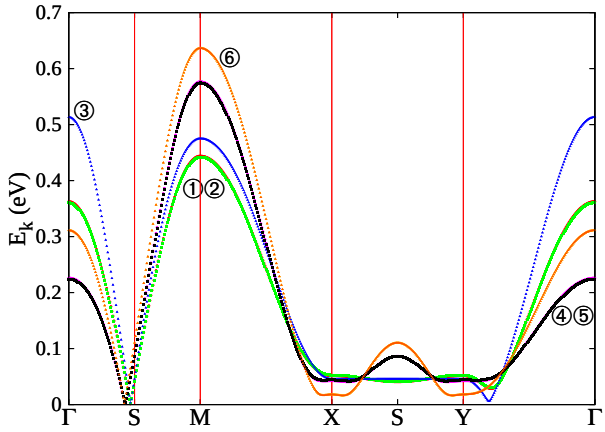

Figure 9: Dispersion relations along the main symmetry lines in the Brillouin zone for $x = 0.175$ ($n = 0.825$). The various curves are explained in main text. Note, the curves 1 and 2, as well as 4 and 5 are practically indistinguishable. This means that the role of the term $\sim c_1$ in W_J (9.209) is negligible.

there is no visible difference between cases 1 and 2, as well as between 4 and 5, differences between cases 1, 3, 4, and 6 are rather pronounced.

9.2.2 Doping dependence of mean-field quantities

Superconducting order parameter. Next, we discuss doping -dependence of the renormalized SC order parameter $\langle \hat{\Delta}_{ij} \rangle_C \equiv \Delta_C$. In Fig. 10 we plot Δ_C for the cases 1-6 specified above as well as for $t_3/t_1 = t'/t = -0.27$ (value being reasonable for BSCCO compounds, [214]), and $J/|t| = 0.3$ (curve 7). Note, that the upper critical concentration x_c for the cases 4 and 5 is close to the VMC result [48], obtained within the Hubbard model for the corresponding values of the model parameters, and by using $|\tilde{\psi}\rangle$ (7.142) instead of $|\psi\rangle$ (7.141).

The presented numerical results (cf. also Table V and Fig. 9) confirm the qualitative predictions made in Sub-subsection 9.1.3. First, nonzero t' enhances superconductivity, in agreement with previous VMC results [50] and other calculations [214]. On the other hand, $\hat{H}_3$ term acts in the opposite direction. The vertical line roughly marks the boundary between under- and over-doped regimes. Importantly, for $|t|/J = 3$, $t'' = 0$ and different t'/t values, x_c lies in the interval $0.2 \lesssim x_c \lesssim 0.35$, depending on the form of $\hat{H}_{tJ}$, as illustrated in Fig. 10. Those results for x_c are in a good overall agreement with the experimental data for the cuprates, [8, 14]. The small difference between the curves 4 and 5 (as well as between 1 and 2) is that, that in the former cases $c_1 = 0$, and shows an insignificant role of the term $\sim c_1$. Note, also, that differences between cases 1-6 become more pronounced with the increasing the doping x.

Superconducting gap and Fermi velocity. In Fig. 11 we plot the doping dependence of the SC gap $D_{\mathbf{k}}$ (9.214) for $\mathbf{k} = (\pi, 0)$, and compare our results with the experimental data [90, 93, 89]. For the selections of t'/t and $J/|t|$ as in Fig. 9, no fully satisfactory agreement with experiment is achieved in the entire range of x. However, the agreement with experiment is reasonable for the parameters corresponding to the curves 1, 2, 4 and 5 in the overdoped regime, both for $|t| = 0.3$ eV and $|t| = 0.4$ eV. The overall agreement is also reasonable for the parameters leading to curve 7. Note, that in all the cases 1-7 the quasiparticle energies obtained here are decisively lower than those in the standard RMFT formulation (c.f. Ref. [61]). These differences are caused by both the particular selection of the renormalization scheme, as well as

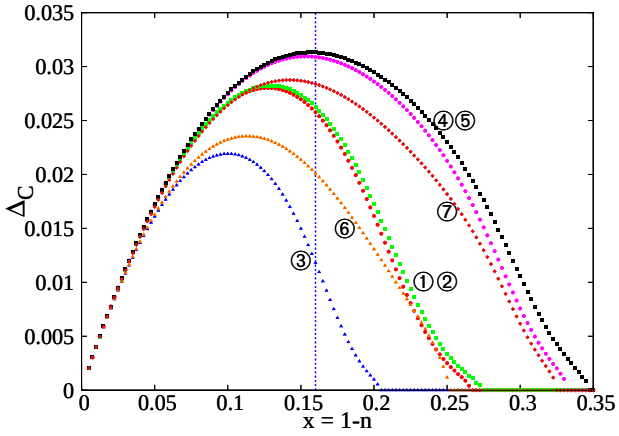

Figure 10: Doping dependence of the renormalized superconducting order parameter $\langle\hat{\Delta}\rangle_C \equiv \Delta_C$. The curves 1-6 correspond to those in Fig. 1. The curve 7 is for $c_1 = 1$, $c_2 = 0$, $t'/t = -0.27$, $t'' = 0$, and for $J/|t| = 0.3$.

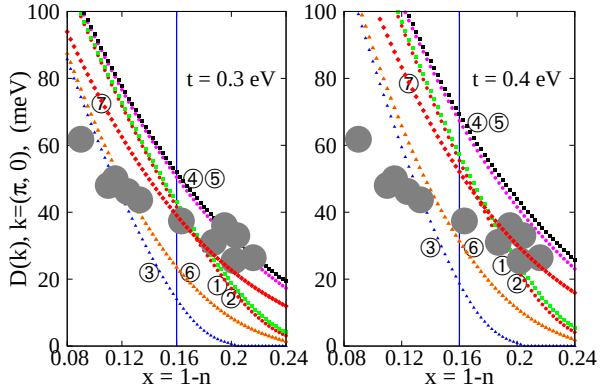

Figure 11: Doping dependences of the SC gap $D_{\mathbf{k}}$ at $\mathbf{k} = (\pi, 0)$ for cases 1-6 and for $t'/t = -0.27$, and $J/|t| = 0.3$ (filled diamonds). Large filled circles - experimental data [88]. Note, that in contradistinction to Ref. [61] no *ad hoc* introduced scaling factor $\alpha = 1/2$ is necessary to obtain a reasonable agreement in the overdoped regime, i.e., to the right of the vertical line.

by the variational method we use. As a consequence, we obtain also lower values of the Fermi velocity

$$v_F = |\nabla_k \xi_{\mathbf{k}}|_{|\mathbf{k}|=k_F} = \sqrt{2}\big(\eta_1 \sin(k_F) + \eta_3 \cos(k_F) + 2\eta_5 \sin(k_F)\big), \tag{9.220}$$

calculated for the nodal ($(0,0) \rightarrow (\pi,\pi)$) direction. The lattice constant has been taken as $a = a_0 = 4$Å [61]. The x-dependence of v_F is detailed in Fig. 12 for the same set of parameters as in Figs. 10 and 11, for both $|t| = 0.3$ eV and $|t| = 0.4$ eV, and compared with the data discussed before [61]. The theoretical values are still too low. Also, the x-dependence of both $D_{\mathbf{k}=(\pi,0)}$ and v_F, obtained within the MF approaches, is stronger than observed in experiment. This feature is shared with the other mean-field approaches [48, 61]. However, the experimental values for BSCCO $\sim$ 1.5 - 1.6 eV Å have also been reported [89, 93], and are quite close to our results.

Influence of the third-nearest hopping integral t''. So far, we have assumed that $t_5 = t'' = 0$. Now we are going to analyze the influence of a nonzero t'' on the doping dependence of the superconducting order parameter $\langle\hat{\Delta}\rangle_C \equiv \Delta_C$. In particular, it is interesting to see

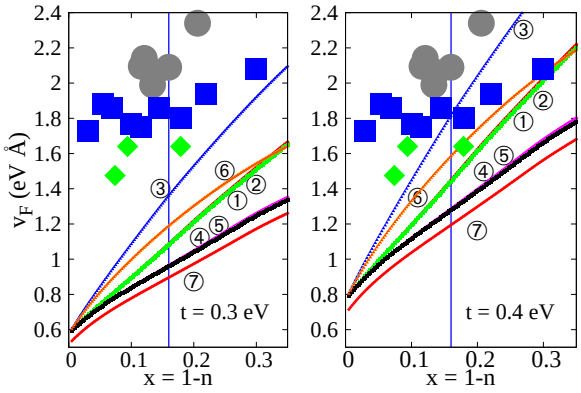

Figure 12: Doping dependence of Fermi velocity in the nodal ($(0,0) \to (\pi,\pi)$) direction. Experimental data (cf. [61] and References therein) are marked by diamonds (YBCO), squares (LSCO) and solid circles (BSCCO).

how $t'' \neq 0$ affects value of the upper critical concentration x_c. We may expect, that the role of t'' is nontrivial, because this parameter determines geometry, or even topology of the Fermi surface [50, 87]. Therefore, it may affect a delicate balance between the normal and the superconducting states.

In Fig. 13 we plot the renormalized superconducting gap Δ_C as a function of doping for $c_1 = 1$, $c_2 = 0$, $t'/t = -0.27$, $J/|t| = 0.3$, and for different values of t'', $0.0 \leq |t''| \leq 0.5|t'|$, with $t''t' < 0$. We observe quite strong t''-dependence of $\Delta_C(x)$. In particular, x_c changes in a non-

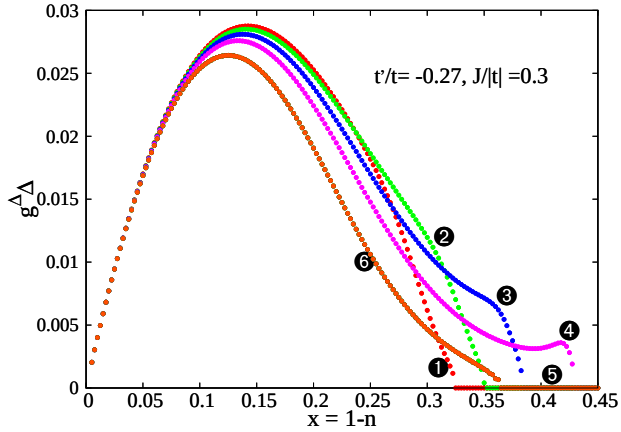

Figure 13: Doping dependence of the renormalized superconducting order parameter Δ_C for $c_1 = 1$, $c_2 = 0$, $t'/t = -0.27$, $J/|t| = 0.3$, and for different values of t''. The curves 1-6 correspond to $-t''/t' = 0.0, 0.1, 0.2, 0.3, 0.4$ and 0.5, respectively.

monotonic manner in the range $0.43 \gtrsim x_c \gtrsim 0.32$. Parenthetically, essentially the same picture is obtained when the hopping to the next and next-next- nearest neighbors are renormalized in the simplest way, i.e., with either y_3 or y_5 in W_t (9.208) equal zero. However, our former conclusion remains valid. Namely, the value of the upper critical concentration remains in a reasonable agreement with the experimental data for the cuprates.

9.2.3 Dependence of upper critical concentration on value of exchange integral J

To provide an additional support for our conclusions, in Table VI we list the values of x_c as a function of J, for either $t'/t = -0.1$ (considered to be relevant to the LSCO compound, [8]) or $t'/t = -0.27$ (value that seems to be more relevant for BSCCO compound).[56] Let us note, that the J- dependence of x_c is substantial; critical concentration is roughly proportional to J.

Table VI. Upper critical concentration x_c vs. exchange integral J, for either $t' = -0.27t$ or $t' = -0.1t$. The symbols A (B) label the cases $c_2 = 0$ and $c_1 = 0$ ($c_1 = 1$), respectively, whereas C means that $c_1 = c_2 = 1$.

$J/\|t\|$	0.2	0.3	0.333	0.375	0.4
$t'/t = -0.1$ A	0.18	0.26	0.29	0.32	0.33
$t'/t = -0.27$ A	0.2	0.31	0.34	0.38	0.4
$t'/t = -0.1$ B	0.18	0.27	0.3	0.33	0.35
$t'/t = -0.27$ B	0.2	0.33	0.36	0.4	0.42
$t'/t = -0.1$ C	0.15	0.21	0.22	0.24	0.25
$t'/t = -0.27$ C	0.15	0.23	0.26	0.28	0.29

[56] Such selection of the values of t'/t ratio may seem to be rather arbitrary. The problem of choice of the values of model parameters for the specific compound is very nontrivial, and will not be discussed in the present Thesis.

10 Summary and discussion

10.1 Summary

The aim of this Thesis has been twofold. Our original aim was to provide a description of the selected properties of the cuprate superconductors within the renormalized mean-field theory (RMFT) of the t-J model. This topic was discussed in detail in part III. As a 'side subject' of this investigation, we have developed a new method of approach, which is based on the maximum entropy (MaxEnt) principle, and which allows to treat MF models considered here in a way which is both internally consistent and optimal from the point of view of statistical physics.

The method is presented in detail in Part II of the Thesis. The formal results presented there are of general validity. Namely, our approach is applicable to any lattice fermion or spin model, and generalization to the case of lattice boson systems is also possible in principle. Because Part II has got its own Summary (Section 5), therefore below we summarize and overview only the contents of Part III.

10.1.1 Comparison of present variational method and non-variational approach based on self-consistent equations

The formal method developed in Part II has been applied in Part III to the RMFT of the t-J model. First, on the example of the mean-field treatment of the simplest form of the t-J Hamiltonian, we have compared two distinct ways of solving mean-field models: the present variational (*var*) method, and the non-variational approach based on the Bogoliubov-de Gennes (BdG) self-consistent (*s-c*) equations. The latter approach to RMFT has been used by many Authors in their studies of various symmetry-broken states encountered in the cuprates [63, 73, 74, 75, 78] and elsewhere. Our results show clearly, that there are nontrivial quantitative, and sometimes even qualitative differences between the *var* and the *s-c* approach; therefore, those two methods cannot be regarded as equivalent. Together with some purely theoretical arguments in favor of the *var* approach, discussed in Part II of the Thesis, this observation leads us to the following conclusion: the *s-c* method should be abandoned in the future analysis of strongly correlated systems, and its existing results should be re-examined within the present *var* approach.

10.1.2 Comparison of different renormalization schemes

We have compared various RMFT Hamiltonians, differing by the way in which correlated averages are approximated, i.e., differing by the applied *renormalization schemes* (RS). It turns out, that differences between various RS are rather pronounced, for both the *var*, and for the *s-c* method. Therefore, also the energy difference between various symmetry-broken states may strongly depend on the particular choice of RMFT scheme. Because the value of the ground state energy (or the free energy) usually serves as the criterion for selecting the optimal mean-field state, the (in)stability of complex states such as the stripe phases [73, 74, 75] or the valence-bond solid phase [78] may be an artifact of a particular version of RS that has been used. This scenario seems to be even more likely due to the circumstance, that the energy differences between the latter states and some reference state (i.e., homogeneous RVB solution) are usually small.

10.1.3 Optimal renormalization scheme

Eventually, we have selected the renormalization scheme (RS) of Fukushima [76], which, in our opinion is the optimal one among RS considered in the present Thesis. Our choice has been motivated by a transparent way in which this RS is constructed, and the fact that the formalism of Ref. [76] may be relatively easily applied to an arbitrary version of the t-J Hamiltonian, e.g. the complete form of the latter, in which the three-site terms are included. Also, the agreement of the results obtained previously by Fukushima with those of the VMC method speak in favor of this particular RS.

Next, the RS of Ref. [76] has been applied to investigate various forms of the t-J Hamiltonian, and for different values of model parameters, i.e., t, t', t'', and J/t. Our results agree quite well with those of the earlier VMC studies (some obtained for the Hubbard model but for the corresponding values of the model parameters, cf. e.g. [48]). For example, doping dependence of the superconducting order parameter, and, in particular, value of the upper critical concentration x_c are close to the VMC results. Also, we have observed the tendency that the nonzero next-nearest neighbor hopping integral t', enhances superconductivity, whereas the three-site term of the t-J Hamiltonian (frequently neglected in the literature) acts in the opposite direction. The influence of the third-nearest neighbor hopping t'' is more complicated; $t'' \neq 0$ enhances superconductivity as compared to the $t'' = 0$ case. However, x_c depends then on t'' in a non-monotonic manner. Also, selected spectral features, e.g. the superconducting gap $D_{\mathbf{k}}$ at $\mathbf{k} = (\pi, 0)$ or the Fermi velocity $v_F = |\nabla_k \xi_{\mathbf{k}}|_{|\mathbf{k}|=k_F}$ in the nodal $((0,0) \rightarrow (\pi,\pi))$ direction have been studied and found to be in a semiquantitative agreement with the ARPES measurements [90, 91, 93], at least in the overdoped regime.

In general, the results obtained within the RMFT turned out to depend quite strongly on both the various forms of the original t-J model, as well as on choice of the model parameters. The right choice of the values of model parameters is nontrivial; the values selected here are identical or close to those selected frequently in the literature.

10.1.4 Limitations of RMFT approach

It must be emphasized here, that even though the RMFT of the t-J model proved to be quite successful in predicting many physical properties of high-T_c cuprate superconductors, it still certainly cannot be regarded as a complete description of those compounds.

First, let us note, that the choice of the mean-fields which may have non-zero optimal (equilibrium) values, i.e being relevant for the problem at hand, is dictated mostly by the experiment. For example, there is a strong experimental evidence in favor of $d_{x^2-y^2}$ symmetry of the superconducting order parameter in the bulk of the cuprates, and very sparse for other possible, e.g. $d_{x^2-y^2} + id_{xy}$ [12]. Therefore, usually from the very beginning we exclude 'by hand' the more exotic symmetries of the superconducting order parameter. However, from the point of view of a model description, certain symmetry-breaking patterns excluded by the experimental data are certainly legitimate. Obviously, there is no certainty, that the solution suggested by experiment would turn out to be the optimal one (i.e., of the lowest energy) within our MF model for the reasonable choice of the model parameters. In connection with this, it must be emphasized, that one of the possible extensions of the present approach to the case a magnetically-ordered states leads to the results which are unphysical in the context of the physics of cuprates. Namely, recently it has been shown by Kaczmarczyk et al. [216], that within the so-called local spin-dependent version of the formalism of Ref. [76], the normal state with saturated, uniform magnetic polarization has lower energy then any state analyzed in the present Thesis, for the same values of the model parameters. We may still hope, that the application of an alternative 'local spin-independent constraint' of Ref. [76] would lead to different results for the magnetic states. Also, the unphysical predictions may be attributed

to the fact, that the effects of the orbital magnetism has been neglected. The same problem is very likely to be present in other versions of RMFT formalism, i.e., for other versions of Gutzwiller approximation (other renormalization schemes).

RMFT leads to description in terms of an effective, *single-particle* Hamiltonian. This feature allows us to address easily all the single-particle properties (e.g. the Fermi surface existence, quasiparticle dispersion observed in ARPES experiments, or symmetry of the order parameter), but obviously must be regarded as (rather crude) approximation to the complete description of the real, strongly-correlated system. Also, for the normal state, single-particle models lead to the description in the spirit of the Landau theory of Fermi liquids (presence of heavy quasiparticles). The latter theory, however, is regarded legitimate at most in the overdoped regime of the phase diagram of the cuprates, as was presented here.

Moreover, as may be intuitively expected, and as has been formally shown in Part II, RMFT of the t-J model (similarly to any simple MF approach) remains in the paradigm of the Landau (-Ginzburg-Wilson) theory of phase transitions, with all consequences of this fact. For example, RMFT may be incapable of providing a proper description of those quantum states, which have no long-range static order.

Another point is an important role of fluctuations (both quantum and thermal) of the order parameters. This issue is probably crucial for understanding of either the nature of the pseudogap phase (cf. e.g. [218]), or the fate of the superconducting state at low doping, and a disappearance of the antiferromagnetic long-range order with the increasing doping. Still, inclusion of fluctuations on the mean-fields level seems to be a highly nontrivial task. Despite some attempts to treat fluctuation within the present method we have made (see the discussion in Subsection 12.1), resolution of this problem remains unclear for us at the moment.

10.2 Outlook and possible extensions of present work

The approach presented in this Thesis may be generalized and improved in several ways, which are listed below.

10.2.1 Approximation-free evaluation of correlated averages

What seems to be the most important modification of the present formalism, is the implementation of a more accurate way of evaluating the correlated averages $\langle\hat{\mathcal{O}}\rangle_C$ (7.144), i.e., averages of operators calculated with respect to the correlated state $|\Psi_0\rangle = \hat{P}_C|\Psi_0\rangle$ (7.139). Even if an exact evaluation of $\langle\hat{\mathcal{O}}\rangle_C$ may turn out to be impossible in practice, still, we should at least try to estimate the error in a rigorous way. In this manner, it might be possible to obtain an exact upper bound for the ground state energy of the t-J Hamiltonian.

10.2.2 Other trial correlated states

In connection with the statements of the previous Sub-subsection, we should mention, that the RVB state used here (Eq. (7.137)) may be not sufficient as a complete trial variational state for the t-J model. Namely, the correlation operator $\hat{P}_C$ in the correlated trial state (7.139) should belong to the same class as the Hamiltonian, for which this trial state serves as an approximate ground state. By this, we understand the following: in contrast to the Hubbard model, the t-J model contains intersite interaction terms. Consequently, the correlator should also belong to the class of inter-site operators, which is not the case for the local Gutzwiller projector $\hat{P}$ (6.128) [217]. This observation suggests, that within the RMFT approach we should use variational trial states of a more general form, e.g. containing various Jastrow correlators [33, 181]. Such states have been already used in some of the VMC studies [53].

10.2.3 Higher-order corrections to t-J Hamiltonian

As discussed in Section 6.135, t-J Hamiltonian is derived from the Hubbard Hamiltonian in the strong coupling limit ($t \ll U$) by using canonical transformation. When doing this, we retain only terms $\sim t$ (kinetic energy part) and $\sim t \cdot (t/U)$ (interaction part). However, using canonical transformation we may construct effective Hamiltonians, which contain terms of higher order, i.e., $t \cdot (t/U)^k$. On the other hand, for the cuprates the reasonable value of t/U is about $1/12$, hence this parameter is not small enough to *a priori* disregard the higher-order terms in the effective Hamiltonian. Therefore, investigation of such effective Hamiltonians (at least the one containing terms up to $\sim t \cdot (t/U)^3$, e.g. the ring-exchange terms) by means of the appropriate mean-field approach seems to be desirable, although it may turn out cumbersome.

10.2.4 Analysis a nonzero temperature

The next problem, which also seems to be important as a subject of future analysis, is an extension of the RMFT approach to arbitrary temperatures. Such formalism would allow us to study large part of x-T phase diagram of the cuprates, and, in particular, to determine the critical temperature T_c as a function of doping. As mentioned in Section 7.4.1, attempts to extend Gutzwiller approximation (GA) to $T > 0$ have been made in the case of the Hubbard model [209, 210, 211]. Recently, a similar approach has been proposed for the t-J-U model [202], hence it is applicable also to the RMFT of the t-J model.

At first glance, it may seem that by using mean-field approach developed in this Thesis, it is legitimate to study RMFT at arbitrary value of T. Unfortunately, this is not the case. Although our method is an intrinsically finite-temperature approach, in the case of RMFT of the t-J model it is not expected to lead to physically reliable predictions at higher temperatures. This insufficiency is not caused by the flaws of the method, but rather by the nature and origin of the RMFT approach. The latter, as based on the Gutzwiller approximation, has been originally devised to study the ground-state properties of the system. Therefore, in this Thesis, we have addressed only the low temperature case, $k_B T/J = 0.002$, which is practically equivalent to the true $T = 0$ situation.

In order to treat consistently arbitrary temperatures, the original method of Gutzwiller should be modified. In our opinion, this problem has no satisfactory solution as yet, at least in the case of the fully-projected trial states required for the t-J model. In particular, in the latter case the formalism of Ref. [202] does not provide a acceptable solution, as discussed in [83].

Nonetheless, because our method is a natural and convenient formal basis for a finite-temperature mean-field analysis of any model of correlated lattice fermions, we hope that it may be also used to construct an appropriate finite-temperature version of RMFT. We should see the progress along this line in the near future.

10.2.5 More complex symmetry-broken states, lattice geometry, and band structure

Another natural path to extend the present treatment would be to investigate a more complex lattice geometry and band structure. For example, bilayer structure and orthorhombic distortion (lack of the C_{4v}-symmetry of CuO_2 planes, leading e.g. to the mixed d- and s wave symmetry of the superconducting order parameter) may be quite easily included within the present formalism. Both features may be important for more realistic description of the cuprate superconductors.

Moreover, the 'classical phonons', i.e., static lattice distortion [199, 204] may be implemented in a simple way within the present approach. This step may be important for the analysis of

the C_{4v}-symmetry breaking, not only in the presence of the orthorhombic distortion, but also spontaneous, as occurring for the x-y symmetric Hamiltonian in the normal state (Pomeranchuk instability, PI).

Next, apart from the homogeneous, non magnetic resonating-valence bond RVB superconducting state, C_{4v}-symmetry broken normal PI state, and simple staggered-flux state (SF) discussed in the present Thesis, other symmetry-broken states may be also studied. In particular, the antiferromagnetic long-range order (AF), its competition or coexistence with both the superconducting, staggered-flux or charge order are among those physical properties, which can be relatively easily described by the present MF approach. Work along these lines is being carried out in our group at present and will be reported separately.

Finally, let us note, that the investigation of yet more complicated states, like the stripe phases or valence bond solid phase [78] within the formalism presented in this Thesis should lead to quantitatively or even qualitatively new results. Also, no attempt to describe the pseudogap state has been made in the present Thesis.

10.2.6 Ginzburg-Landau potential

Using the results of Section 4.5, we can construct the Ginzburg-Landau (G-L) free-energy potentials (i.e., a discrete version of the original G-L functional) for the RMFT Hamiltonians. The Ginzburg-Landau picture for the RMFT approach may be helpful to study, or at least to visualize, the normal-to-superconducting transition at x_c. It would be interesting then to compare the Ginzburg-Landau approach resulting from the microscopic RMFT description with the phenomenological G-L theory as given e.g. in Ref. [219].

Part IV
Appendices and supplementary material

11 Appendices

11.1 Appendix A: Deficiencies of approach not based on the method of Lagrange multipliers

In Sub-subsection 4.3.3, we have point out some difficulties arising when one tries to apply the MaxEnt principle to the MF approach, but without invoking the method of Lagrange multipliers. For completeness, below we present further comments on that issue.

At first, let us *assume*, that the eigenbasis of the (yet unknown) MF density operator $\hat{\rho}$, i.e., $\{|i\rangle\}_{i=1}^{D_H}$, is $\vec{A}$-independent. Using $\{|i\rangle\}_{i=1}^{D_H}$, $\mathcal{S}_\lambda$ (4.9) can be written as

$$\mathcal{S} = -\sum_i \Big(p_i \ln p_i + \beta [\hat{K}(\vec{A})]_{ii} p_i + \omega (p_i - 1/D_H) \Big) , \tag{11.221}$$

with $[\hat{O}]_{ii} \equiv \langle i|\hat{O}|i\rangle$ and $[\hat{K}(\vec{A})]_{ii} = [\hat{H}(\vec{A})]_{ii} - \mu[\hat{N}]_{ii}$. In the present situation, Eqs. (4.12), i.e.,

$$\mathrm{Tr}[\hat{\rho}\hat{A}_s] = \langle \hat{A}_s \rangle = A_s = \sum_i p_i [\hat{A}_s]_{ii} = \sum_i p_i \langle i|\hat{A}_s|i\rangle, \tag{11.222}$$

allows us to eliminate A_s in favor of the probabilities p_i, and the latter remain then the only independent variables. Consequently, the necessary conditions for $\mathcal{S}$ to have an extremum read now

$$\frac{\partial \mathcal{S}}{\partial p_k} = 0 = -\ln p_k - \beta [\hat{K}(\vec{A})]_{kk} - \omega - 1 - \beta \sum_i p_i \sum_s \frac{[\hat{K}(\vec{A})]_{ii}}{\partial A_s} [\hat{A}_s]_{kk}, \tag{11.223}$$

for $k = 1, 2, \ldots, D$ and with the normalization condition ($\sum_i p_i = 1$) imposed. However, equations (11.223) cannot be solved analytically, and thus an explicit functional dependence of p_i on T, V (or Λ), $\mu, \vec{h}, \ldots$ cannot be obtained. Eqs. (11.223) can be rewritten as

$$p_k = \mathcal{Z}^{-1} \exp \left(-\beta ([\hat{K}(\vec{A})]_{kk} + \sum_i p_i \sum_s \frac{\partial [\hat{K}(\vec{A})]_{ii}}{\partial A_s} [\hat{A}_s]_{kk} \right) . \tag{11.224}$$

Obviously, the probability distribution given by (11.224) is not of a canonical (Gibbs) form. By defining

$$-\bar{\lambda}_s = \left\langle \frac{\hat{K}(\vec{A})}{\partial A_s} \right\rangle = \sum_i p_i \frac{[\hat{K}(\vec{A})]_{ii}}{\partial A_s}, \tag{11.225}$$

we can rewrite (11.224) as

$$p_k = \mathcal{Z}^{-1} \exp \big(-\beta ([\hat{K}(\vec{A})]_{kk} - \sum_s \bar{\lambda}_s [\hat{A}_s]_{kk} \big). \tag{11.226}$$

For a given $\vec{A}$, p_k (11.226) has a form similar to that of q_k, given by diagonal elements of $\hat{\rho}_\lambda$ (4.29), i.e.,

$$q_k = \mathcal{Z}_\lambda^{-1} \exp \Big(-\beta \big([\hat{K}(\vec{A})]_{kk} - \sum_s \lambda_s ([\hat{A}_s]_{kk} - A_s) \big) \Big), \tag{11.227}$$

because the constant (independent of k) term present in q_k (11.227) and absent in p_k (11.226) can be absorbed in $\mathcal{Z}_\lambda^{-1}$.

Apparently, by defining (11.225), we have taken a step towards approach of Subsections 4.3 and 4.4. However, first, a route presented here ignores the algebraic structure of the quantum model. It should be *a posteriori* checked, if indeed the assumption we have made, i.e., the existence of an $\vec{A}$ -independent eigenbasis of $\hat{\rho}$ (which has been used to compute trace appearing in the above formulas) does not lead to contradiction. This task, however, is enormously complicated, if not entirely impossible. On the other hand, we cannot assume *a priori*, that MF grand Hamiltonian $\hat{K}(\vec{A})$ commutes with the density matrix, therefore the eigenbasis of the former operator cannot serve as the eigenbasis of the latter.

If the eigenbasis of $\hat{\rho}$ is not $\vec{A}$-independent, the reasoning leading to most of the above formulas ceases to valid. To see that, let us expand eigenvectors of $\hat{\rho}$ in some $\vec{A}$-independent basis, i.e., $|i(\vec{A})\rangle = \sum_l c_{il}(\vec{A})|l\rangle$ and next rewrite (11.222) in a form

$$A_s = \sum_i p_i[\hat{A}_s]_{ii} = \sum_i p_i \langle i(\vec{A})|\hat{A}_s|i(\vec{A})\rangle = \sum_i p_i \sum_{j,l} c^*_{ij}(\vec{A}) c_{il}(\vec{A}) \langle j|\hat{A}_s|l\rangle. \tag{11.228}$$

Apart from very special cases, A_s variables cannot be expressed as any explicitly known functions of p_i.

11.2 Appendix B: Equivalence of two alternative expressions for the second derivative of thermodynamic grand potential Ω

Below we show the equivalence of two alternative formulas (Eq. (4.52) and Eq. (4.53)) for the second derivative of the thermodynamic potential Ω with respect to thermodynamic variables $x_1, x_2 \in \{T, V, \mu, \vec{h}, \ldots\}$. Eqs. (4.52) and (4.53) read, respectively

$$\frac{\partial^2\Omega}{\partial x_2 \partial x_1} = \frac{\partial}{\partial x_2}\left(\frac{\partial \mathcal{F}}{\partial x_1}\right)_0 = \left(\frac{\partial^2 \mathcal{F}}{\partial x_2 \partial x_1}\right)_0 + \sum_s \left\{ \left(\frac{\partial^2 \mathcal{F}}{\partial x_1 \partial A_s}\right)_0 \frac{\partial A_{s0}}{\partial x_2} + \left(\frac{\partial^2 \mathcal{F}}{\partial x_1 \partial \lambda_s}\right)_0 \frac{\partial \lambda_{s0}}{\partial x_2} \right\} \tag{11.229}$$

$$\begin{aligned}
\frac{\partial^2\Omega}{\partial x_2 \partial x_1} &= \frac{\partial}{\partial x_2}\left[\left(\frac{\partial \mathcal{F}}{\partial x_1}\right)_0 + \sum_s \left\{ \left(\frac{\partial \mathcal{F}}{\partial A_s}\right)_0 \frac{\partial A_{s0}}{\partial x_1} + \left(\frac{\partial \mathcal{F}}{\partial \lambda_s}\right)_0 \frac{\partial \lambda_{s0}}{\partial x_1} \right\} \right] \\
&= \left(\frac{\partial^2 \mathcal{F}}{\partial x_1 \partial x_2}\right)_0 + \sum_s \left\{ \left(\frac{\partial^2 \mathcal{F}}{\partial x_1 \partial A_s}\right)_0 \frac{\partial A_{s0}}{\partial x_2} + \left(\frac{\partial^2 \mathcal{F}}{\partial x_1 \partial \lambda_s}\right)_0 \frac{\partial \lambda_{s0}}{\partial x_2} \right\} \\
&+ \sum_s \left\{ \left(\frac{\partial^2 \mathcal{F}}{\partial x_2 \partial A_s}\right)_0 \frac{\partial A_{s0}}{\partial x_1} + \left(\frac{\partial^2 \mathcal{F}}{\partial x_2 \partial \lambda_s}\right)_0 \frac{\partial \lambda_{s0}}{\partial x_1} \right\} \\
&+ \sum_{s,t} \left\{ \left(\frac{\partial^2 \mathcal{F}}{\partial A_t \partial A_s}\right)_0 \frac{\partial A_{s0}}{\partial x_1}\frac{\partial A_{t0}}{\partial x_2} + \left(\frac{\partial^2 \mathcal{F}}{\partial A_s \partial \lambda_t}\right)_0 \frac{\partial A_{s0}}{\partial x_1}\frac{\partial \lambda_{t0}}{\partial x_2} \right. \\
&+ \left. \left(\frac{\partial^2 \mathcal{F}}{\partial A_s \partial \lambda_t}\right)_0 \frac{\partial A_{s0}}{\partial x_2}\frac{\partial \lambda_{t0}}{\partial x_1} + \left(\frac{\partial^2 \mathcal{F}}{\partial \lambda_s \partial \lambda_t}\right)_0 \frac{\partial \lambda_{s0}}{\partial x_1}\frac{\partial \lambda_{t0}}{\partial x_2} \right\} \\
&+ \sum_s \left\{ \left(\frac{\partial \mathcal{F}}{\partial A_s}\right)_0 \frac{\partial^2 A_{s0}}{\partial x_1 \partial x_2} + \left(\frac{\partial \mathcal{F}}{\partial \lambda_s}\right)_0 \frac{\partial^2 \lambda_{s0}}{\partial x_1 \partial x_2} \right\}.
\end{aligned} \tag{11.230}$$

The r.h.s. of (11.229) is asymmetric with respect to x_1 and x_2. It seems also to be incompatible with (11.230). However, both features are only apparent. Indeed, for any s we have

$$0 = \frac{\partial}{\partial x_2}\left(\frac{\partial \mathcal{F}}{\partial A_s}\right)_0 = \left(\frac{\partial^2 \mathcal{F}}{\partial A_s \partial x_2}\right)_0 + \sum_t \left\{ \left(\frac{\partial^2 \mathcal{F}}{\partial A_t \partial A_s}\right)_0 \frac{\partial A_{t0}}{\partial x_2} + \left(\frac{\partial^2 \mathcal{F}}{\partial A_s \partial \lambda_t}\right)_0 \frac{\partial \lambda_{t0}}{\partial x_2} \right\}. \tag{11.231}$$

Multiplying (11.231) by $\partial A_{s0}/\partial x_1$ and summing over s we obtain

$$0 = \sum_s \left(\frac{\partial^2 \mathcal{F}}{\partial A_s \partial x_2}\right)_0 \frac{\partial A_{s0}}{\partial x_1} + \sum_{s,t} \left\{ \left(\frac{\partial^2 \mathcal{F}}{\partial A_t \partial A_s}\right)_0 \frac{\partial A_{s0}}{\partial x_1}\frac{\partial A_{t0}}{\partial x_2} + \left(\frac{\partial^2 \mathcal{F}}{\partial A_s \partial \lambda_t}\right)_0 \frac{\partial A_{s0}}{\partial x_1}\frac{\partial \lambda_{t0}}{\partial x_2} \right\}. \tag{11.232}$$

Analogously,

$$0 = \frac{\partial}{\partial x_2}\left(\frac{\partial \mathcal{F}}{\partial \lambda_s}\right)_0 = \left(\frac{\partial^2 \mathcal{F}}{\partial \lambda_s \partial x_2}\right)_0 + \sum_t \left\{ \left(\frac{\partial^2 \mathcal{F}}{\partial A_t \partial \lambda_s}\right)_0 \frac{\partial A_{t0}}{\partial x_2} + \left(\frac{\partial^2 \mathcal{F}}{\partial \lambda_t \partial \lambda_s}\right)_0 \frac{\partial \lambda_{t0}}{\partial x_2} \right\}. \tag{11.233}$$

Multiplying (11.233) by $\partial \lambda_{s0}/\partial x_1$ and summing over s we obtain

$$0 = \left(\frac{\partial^2 \mathcal{F}}{\partial A_s \partial x_2}\right)_0 \frac{\partial \lambda_{s0}}{\partial x_1} + \sum_{s,t} \left\{ \left(\frac{\partial^2 \mathcal{F}}{\partial A_t \partial \lambda_s}\right)_0 \frac{\partial \lambda_{s0}}{\partial x_1}\frac{\partial A_{t0}}{\partial x_2} + \left(\frac{\partial^2 \mathcal{F}}{\partial \lambda_s \partial \lambda_t}\right)_0 \frac{\partial \lambda_{s0}}{\partial x_1}\frac{\partial \lambda_{t0}}{\partial x_2} \right\}. \tag{11.234}$$

Changing dummy indices (e.g. $s \leftrightarrow t$) if necessary, we identify in (11.230) the terms appearing on the r.h.s. of Eq. (11.232) and (11.234). Note also, that the last line of (11.230) vanishes due to (4.34). This completes the proof of the equivalence of Eq. (11.229) and Eq. (11.230) (or Eq. (4.52) and Eq. (4.53)).

11.3 Appendix C: Generalized thermodynamic potentials and Legendre transformations

Natural thermodynamic variables of $\mathcal{F}_z(\vec{A}) \equiv \mathcal{F}_z(T, V, \mu, \vec{h}; \vec{A})$ (4.44) are identical with those of the thermodynamic grand potential $\Omega = \Omega(T, V, \mu, \vec{h})$. The relation between $\mathcal{F}_z$ and Ω given by Eqs. (4.47), i.e.,

$$\mathcal{F}_z(\vec{A}_0(T, V, \mu, \vec{h}); T, V, \mu, \vec{h}) = \Omega(T, V, \mu, \vec{h}), \tag{11.235}$$

is same as in the original Landau theory. Therefore $\mathcal{F}_z(\vec{A})$ corresponds to $\Omega(\mu, T, \eta)$, Eq. (146.3) of Ref. [148]. Nonetheless, within the original Landau approach, other generalized thermodynamic potentials are also used, e.g. $\Phi(p, T, \eta)$, Eq. (143.1) of [148] ('Gibbs free energy potential'), denoted here as $\mathcal{G}(T, p, N, \vec{h}; \vec{A})$. Because within the phenomenological approach of Landau, there exist no underlying microscopic mean-field model, one has a freedom to choose those thermodynamic variables and the corresponding generalized thermodynamic potential, which are the most convenient (or natural) for the considered problem. On the other hand, within the present approach, $\mathcal{F}_z(T, V, \mu, \vec{h}; \vec{A})$, appearing as the result of the MF grand-canonical description, has a privileged status. Similarly, free energy Landau potential $\mathcal{F}_z(T, V, \vec{h}; \vec{A})$ would be a natural choice for the canonical ensemble, in a situation when the average particle number does not appear as one of the mean fields, e.g. for the MF model of the spin system analyzed in Section 12.3.

If we want to define here the analog of $\Phi(p, T; \eta)$ of Landau, we have to make Legendre transformations $V \leftrightarrow p$, $\mu \leftrightarrow N$ leading from $\{T, V, \mu, \vec{h}\}$ to the new set of independent thermodynamic variables, i.e., $\{T, p, N, \vec{h}\}$. However, in order to do that, we have to construct first the equilibrium thermodynamic potentials along the lines of Subsection 4.6. Both p and N become then functions of T, V, μ, and $\vec{h}$, given by the respective derivatives of $\Omega = \Omega(T, V, \mu, \vec{h})$, cf. Eqs. (4.48). By inverting those relations, we obtain $V(T, p, N, \vec{h})$ and $\mu(T, p, N, \vec{h})$. Consequently, $\mathcal{G}(T, p, N, \vec{h}; \vec{A})$ should be defined as

$$\mathcal{G}(T, p, N, \vec{h}; \vec{A}) = \mathcal{F}_z(T, V(T, p, N, \vec{h}), \mu(T, p, N, \vec{h}), \vec{h}; \vec{A}) + pV + \mu N. \tag{11.236}$$

In the above, N denotes the equilibrium value of the average particle number, and not the corresponding mean field $A_1 = \langle \hat{N} \rangle_z$. In analogy to Eq. (11.235), we have

$$\mathcal{G}(T, p, N, \vec{h}; \vec{A}_0^G(T, p, N, \vec{h})) = G(T, p, N, \vec{h}). \tag{11.237}$$

Superscript G indicates that now $\vec{A}_0^G$ depend on $T, p, N, \vec{h}$, and not on $T, V, \mu, \vec{h}$. Even if within the present approach, $\mathcal{G}(T, p, N, \vec{h}; \vec{A})$ appears in a less natural manner than $\mathcal{F}_z(T, V, \mu, \vec{h}; \vec{A})$, it may be more convenient in a situation, when the average particle number and pressure are fixed. If needed, it may be constructed as discussed above.

11.4 Appendix D: Renormalization scheme of Fukushima

Below we provide a more detailed analysis of the RS proposed by Fukushima in Ref. [76]. Within this formalism, Gutzwiller projection operator (projector) $\hat{P}$ (6.128) is replaced by the following correlation operator (correlator)[57]

$$\hat{P}_C = \hat{P}^{(F)} = \prod_i (\lambda_{i\uparrow}^F)^{\frac{\hat{n}_{i\uparrow}}{2}} (\lambda_{i\downarrow}^F)^{\frac{\hat{n}_{i\downarrow}}{2}} (1 - \hat{n}_{i\uparrow}\hat{n}_{i\downarrow}). \tag{11.238}$$

As a consequence, the correlated trial state (7.140) is replaced by $|\Psi\rangle$ (7.141) of the form

$$|\Psi\rangle = \hat{P}^{(F)}(\lambda_{i\uparrow}^F, \lambda_{i\downarrow}^F)|\Psi_0\rangle = \prod_i (\lambda_{i\uparrow}^F)^{\frac{\hat{n}_{i\uparrow}}{2}} (\lambda_{i\downarrow}^F)^{\frac{\hat{n}_{i\downarrow}}{2}} (1 - \hat{n}_{i\uparrow}\hat{n}_{i\downarrow})|\Psi_0\rangle. \tag{11.239}$$

With the help of $|\Psi\rangle$ (11.239) we define the correlated average (Eq. (7.144)) as

$$\langle \hat{\mathcal{O}} \rangle_C \equiv \frac{\langle \Psi | \hat{\mathcal{O}} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\langle \Psi_0 | \hat{P}_C \hat{\mathcal{O}} \hat{P}_C | \Psi_0 \rangle}{\langle \Psi_0 | \hat{P}_C^2 | \Psi_0 \rangle}. \tag{11.240}$$

$\hat{P}^{(F)}$ differs from $\hat{P}$ (6.128) by presence of the *fugacity factors*[58] $\lambda_{i\sigma}^F$, see also Eqs. (7.141). Such modification of the standard Gutzwiller projector allows us to enforce the following condition: the uncorrelated average number of electrons on each lattice site i and each spin direction σ is equal to the corresponding correlated average, computed within the approximation of Ref. [76] (see below). Explicitly,

$$\langle \hat{n}_{i\sigma} \rangle_C^{app} = \langle \hat{n}_{i\sigma} \rangle. \tag{11.241}$$

There exist few variants of the formalism of Ref. [76]. The condition (11.241) refers to the 'local spin-dependent constraint'. Alternatively, one may choose the 'local spin-independent constraint'. In such case, we replace $\hat{n}_{i\sigma}$ in (11.241) by $\hat{n}_i \equiv \hat{n}_{i\uparrow} + \hat{n}_{i\downarrow}$. Consequently, then we have only one fugacity factor λ_i^F for each site. There exist also the corresponding 'global', i.e., site-independent variants of both 'spin-dependent' and 'spin-independent' constraint, i.e., with $\hat{n}_{i\sigma}$ ($\hat{n}_i$) in (11.241) replaced by $\sum_i \hat{n}_{i\sigma} = \hat{N}_\sigma$ and $\sum_i \hat{n}_i = \hat{N}$, respectively. However, for non-magnetic ($n_{i\uparrow} = n_{i\downarrow}$) and homogeneous ($n_{i\sigma} = n_{j\sigma}$) solutions analyzed in this Thesis, all the above mentioned variants of the RS of Fukushima are equivalent. Therefore, in our case, the condition (11.241) reduces to

$$\langle \hat{N} \rangle_C^{app} = \langle \hat{N} \rangle = N. \tag{11.242}$$

Note, that when $|\Psi_0\rangle$ is an eigenstate of $\hat{N}$, projection does not change N, because $[\hat{N}, \hat{P}] = 0$. However, this is not the case for $|\Psi_0\rangle$ of the BCS form, because the latter is an eigenstate of

[57] All formulas in this Appendix are provided by using our own notation, which differs from that of the original reference [76].

[58] Please not confuse those quantities with the Lagrange multipliers introduced within our formalism.

the Hamiltonian which does not commute with $\hat{N}$, $[\hat{N}, \hat{H}_{\rm BCS}] \neq 0$. Also, without the fugacity factors $\lambda^F_{i\sigma}$ in the projector, due to the approximation used to compute $W(\vec{A})$ (7.167), we would usually have $\langle \hat{N} \rangle^{app}_C \neq \langle \hat{N} \rangle$ even if $|\Psi_0\rangle$ appearing in $|\Psi\rangle = \hat{P}^{(F)}|\Psi_0\rangle$ (11.239) was an eigenstate of $\hat{N}$.[59] Coming back to Eq. (11.241), we can rewrite it in a form

$$\langle \hat{n}_{i\sigma} \rangle_C = \frac{\langle \lambda^F_{i\sigma} \hat{n}_{i\sigma} \big(1 - \hat{n}_{i\bar{\sigma}}\big) \prod_{l \neq i} \big(\hat{P}^{(F)}_l\big)^2 \rangle}{\langle \prod_l \big(\hat{P}^{(F)}_l\big)^2 \rangle} \tag{11.243}$$

Approximate way of computing (11.243) consist of neglecting all inter-site contractions. This yields

$$\langle \hat{n}_{i\sigma} \rangle_C \approx \langle \hat{n}_{i\sigma} \rangle^{app}_C = \frac{\lambda^F_{i\sigma} \big(1 - n_{i\bar{\sigma}}\big)}{\Xi_i} \cdot n_{i\sigma}, \tag{11.244}$$

where

$$\Xi_i = \big\langle \big(\hat{P}^{(F)}_i\big)^2 \big\rangle = (1 - n_{i\uparrow})(1 - n_{i\downarrow}) + \lambda^F_{i\uparrow} n_{i\uparrow}(1 - n_{i\downarrow}) + \lambda^F_{i\downarrow} n_{i\downarrow}(1 - n_{i\uparrow}). \tag{11.245}$$

By solving (11.244) and (11.245) we obtain

$$\lambda^F_{i\sigma} = \frac{1 - n_{i\sigma}}{1 - n_i}, \qquad \Xi_i = \frac{(1 - n_{i\uparrow})(1 - n_{i\downarrow})}{1 - n_i}. \tag{11.246}$$

Now, using the above results, one can compute various correlated averages (11.240). Note, that the method of Ref. [76] can be relatively easily applied to the full form of the t-J Hamiltonian (6.126), as well as to the more complicated symmetry-broken states. Here, we will not provide details of this procedure, the Reader is encouraged to consult the original paper. Formula for $\langle \hat{H}_3 \rangle^{app}_C$, not appearing in the latter paper, is too lengthy to be reproduced here, but for the special (homogeneous, nonmagnetic) case it is provided in Section 9 (Eq. (9.210)). Also, the average of the full kinetic exchange term, Eq. (9.209) is a special case of the general expression for $\langle \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j - \frac{1}{4} \hat{n}_i \hat{n}_j \rangle^{app}_C$. As an example, we invoke only the formula for the renormalization hopping amplitude

$$\langle c^\dagger_{i\sigma} c_{j\sigma} \rangle^{app}_C = \sqrt{\frac{1 - n_i}{1 - n_{i\sigma}}} \sqrt{\frac{1 - n_j}{1 - n_{j\sigma}}} \left(\chi_{ij\sigma} - \chi_{ij\bar{\sigma}} \frac{\chi_{ij\sigma} \chi_{ij\bar{\sigma}} + \Delta_{ji} \Delta_{ij}}{(1 - n_{i\bar{\sigma}})(1 - n_{j\bar{\sigma}})} \right), \tag{11.247}$$

provided already in Subsection 7.5. Please note, that all formulas in Ref. [76] are provided up to the second order in the inter-site contractions χ^*_{ij}, Δ^*_{ij}, and the estimated corrections are of order of $|\chi_{ij\sigma}|^4$ and $|\Delta_{ij}|^4$. Therefore, all the conditions of the form (11.241) are fulfilled to the same accuracy. However, the results of Ref. [76] may be, at least in principle, systematically extended beyond the second order in χ_{ij} and Δ_{ij}.

[59] Let us note here, that the condition (11.242) can be also enforced in an alternative manner, cf. [220, 221] and References therein.

12 Supplements

12.1 Supplement A: Thermodynamic fluctuations and internal limitations of mean-field description

Below we provide the Reader with the analysis of thermodynamic fluctuations. This issue has been omitted in the main text, as being weakly related to the main scope of the present Thesis. The problem is also quite nontrivial, due to the fact, the within the present approach two kinds of fluctuations are present. On the other hand, even incomplete analysis of the non-equilibrium situation allows us to establish natural internal limitations on the validity of the mean-field approach.

12.1.1 Probability of a non-equilibrium MF configuration

Following Landau [148], (see also [133]), we assume that the probability that our system remains in a state characterized by a mean-field configuration $\vec{A}$, for a given T, V, μ, and $\vec{h}$ is given by

$$w(T,V,\mu,\vec{h};\vec{A}) \equiv w(\vec{A}) \sim \exp\Big(-\beta\Delta\Omega_L(T,V,\mu,\vec{h};\vec{A})\Big). \tag{12.248}$$

After Landau, by $\Delta\Omega_L(T,V,\mu;\vec{A})$ we denote minimum work required to bring the system from $\vec{A}_0$ to $\vec{A}$ in a reversible manner and assuming constant T, V, μ and external field(s) $\vec{h}$. In what follows, it we will be shown that

$$\Delta\Omega_L(T,V,\mu,\vec{h};\vec{A}) \equiv \mathcal{F}_z(T,V,\mu,\vec{h};\vec{A}) - \mathcal{F}_z(T,V,\mu,\vec{h};\vec{A}_0), \tag{12.249}$$

and, consequently that

$$w(\vec{A}) \sim \exp\big(-\beta\mathcal{F}_z(T,V,\mu,\vec{h};\vec{A})\big). \tag{12.250}$$

In order to that, we make use of the results of Ref. [222].[60] It will be shown, that $\mathcal{F}_z(\vec{A})$ may be identified with the 'non-equilibrium grand potential', $\Omega_n(\vec{A})$. To define the latter quantity, we start from a stable equilibrium situation characterized by $\vec{A} = \vec{A}_0$ and fixed T, V, μ and $\vec{h}$. If we want to deal with a reversible process, i.e., the sequence of equilibrium states, we must introduce additional, external forces. On the technical level, it may be realized by introducing additional 'source' terms of the form $(-\sum_s \zeta_s \hat{A}_s)$, which are added to the MF Hamiltonian $\hat{H}_z(\vec{A})$ (4.42). We define also

$$\Psi(\vec{A}) \equiv -\beta^{-1}\ln \mathrm{Tr}\big[\exp\big(-\beta(\hat{H}_z(\vec{A}) - \sum_s \zeta_s\hat{A}_s - \mu\hat{N})\big)\big]. \tag{12.251}$$

For each $\vec{A}$, values of the 'source fields' $\zeta_s = \zeta_s(\vec{A})$ will be determined from the condition, that $\vec{A}$ is an equilibrium mean-field configuration in the presence of those additional forces, i.e., that $\nabla_A\Psi(\vec{A}) = 0$. Then $\Psi(\vec{A})$ (12.251) becomes a corresponding equilibrium grand potential for $\vec{A}$. Next, following Ref. [222], we define the non-equilibrium grand potential $\Omega_n(\vec{A})$ as

$$\Omega_n(\vec{A}) \equiv \Psi(\vec{A}) - U_\zeta = \Psi(\vec{A}) + \sum_s \zeta_s A_s. \tag{12.252}$$

In the above, $U_\zeta = \langle -\sum_s \zeta_s\hat{A}_s\rangle_z = -\sum_s \zeta_s A_s$ is the potential energy of the additional force fields. To show that indeed $\mathcal{F}_z(\vec{A})$ is identical to $\Omega_n(\vec{A})$, we have to invoke the results of

[60] Although in Ref. [222] the non-mean-field case is analyzed, and the free energy is used instead of a grand potential, we follow closely the original reasoning of Leontovich and adapt it for the present purposes.

Subsection 4.8. Namely, after the self-consistency equations (4.41) are solved for each $\vec{A}$, MF Hamiltonians $\hat{H}_z(\vec{A})$ and $\hat{H}_z(\vec{A}) - \sum_s \zeta_s(\vec{A})\hat{A}_s$ belong to the same universality class, and therefore differ only by a constant term $\sum_s \zeta_s(\vec{A})A_s$. Indeed, in such situation, for each $\vec{A}$, value of λ_s appearing in $\hat{H}_z(\vec{A})$ is equal to the value of $(\lambda_s+\zeta_s \equiv \tilde{\lambda}_s^{\zeta})$ appearing in $\hat{H}_z(\vec{A})-\sum_s \zeta_s(\vec{A})\hat{A}_s$. We obtain a desired result due to obvious relation

$$\hat{H}_z(\vec{A}) - \sum_s \zeta_s(\vec{A})\hat{A}_s = \hat{H}(\vec{A}) - \sum_s \tilde{\lambda}_s^{\zeta}(\vec{A})\big(\hat{A}_s - A_s\big) + \sum_s \zeta_s(\vec{A})A_s. \tag{12.253}$$

Consequently, for each $\vec{A}$, also the corresponding functionals $\mathcal{F}_z(\vec{A})$ (4.44) and $\Psi(\vec{A})$ (12.251) differ only by the term $\sum_s \zeta_s A_s$, i.e., $\Psi(\vec{A}) = \mathcal{F}_z(\vec{A}) - \sum_s \zeta_s A_s$. Therefore, from (12.252) it follows that $\mathcal{F}_z(\vec{A}) = \Omega_n(\vec{A})$ and $\Omega_n(\vec{A}) - \mathcal{F}_z(T, V, \mu; \vec{A}_0)$ is equal to $\Delta\Omega_L(T, V, \mu; \vec{A})$, as in [148]. This proves Eq. (12.249). In that way we have argued, that $\mathcal{F}_z(\vec{A})$, defined for a given microscopic mean-field model according to Eq. (4.44), may be consistently interpreted as a Landau potential.

12.1.2 'Classical' and 'quantum' probability distributions

Invoking (12.248) or (12.250), we may define

$$p(\vec{A} \leq \vec{\mathcal{A}} \leq \vec{A} + d\vec{A}) \equiv \varphi(\vec{A})d\vec{A}, \qquad \varphi(\vec{A}) = \mathcal{N}^{-1}\exp(-\beta\mathcal{F}_z(\vec{A})). \tag{12.254}$$

$\vec{\mathcal{A}}$ is a random variable associated with $\vec{A}$, and $\mathcal{N}$ is the normalization constant. In result, we deal with two entirely different probability distributions. The first is given for each $\vec{A}$ by $\hat{\rho}_z(\vec{A})$. It yields a (conditional) probability of finding a system in a quantum state $|i\rangle$ *provided* that values of the classical parameters are equal $\vec{A}$ (therefore, strictly speaking, we should write $|i(\vec{A})\rangle$ instead of $|i\rangle$). The second probability distribution is $\varphi(\vec{A})$ defined by (12.254), and related to the probability of a particular value of each of the 'classical' quantities $A_1, \ldots, A_M$. The latter variables may be regarded as constituting some kind of a 'classical phase space'.

12.1.3 Degenerate minima of $\mathcal{F}_z(\vec{A})$

In the absence of an external, symmetry breaking field $\vec{h}$ (i.e., real magnetic field), $\varphi(\vec{A})$ (12.254) is invariant with respect to the allowed symmetry transformations of the mean fields (order parameters), exactly as in the original Landau approach. However, in the present case the symmetry of $\varphi(\vec{A})$ is inherited from microscopic MF Hamiltonian $\hat{H}_z(\vec{A})$. Namely, two microstates, $|i(\vec{A}')\rangle$ and $|i(\vec{A}'')\rangle$ with $\vec{A}'$ and $\vec{A}''$ related by symmetry transformations (e.g. with A_s for $\vec{A}'$ and $-A_s$ for $\vec{A}''$ if $\mathcal{F}_z(\vec{A})$ is even with respect to A_s), enter the corresponding partition functions $\mathcal{Z}_z(\vec{A}')$ and $\mathcal{Z}_z(\vec{A}'')$(4.42) with exactly the same weight. Consequently, we have $\mathcal{Z}_z(\vec{A}') = \mathcal{Z}_z(\vec{A}'')$ and $\mathcal{F}_z(\vec{A}') = \mathcal{F}_z(\vec{A}'')$. This may be also interpreted from the point of view of MaxEnt inference. Namely, if $\hat{H}_z(\vec{A}') = \hat{H}_z(\vec{A}'')$, no prior information about the symmetry breaking is encoded in $\mathcal{S}_\lambda$, hence $\vec{A}'$ is equivalent to $\vec{A}''$ and the resulting probability distributions $p_1(\vec{A}'), \ldots, p_{D_H}(\vec{A}')$ and $p_1(\vec{A}''), \ldots, p_{D_H}(\vec{A}'')$ are identical. As a consequence, the 'classical' average values of some of the mean-fields, which could be defined as

$$\overline{A_s} \equiv \int_{\mathcal{D}_\mathcal{A}} A_s \varphi(\vec{A}) d\vec{A} \tag{12.255}$$

obtained using $\varphi(\vec{A})$ are equal zero. This feature is obviously connected to the fact, that the symmetry should be explicitly broken *before* any *classical* fluctuations around the equilibrium situation can be considered - classical situation corresponds to already broken symmetry.

Therefore, we have to replace $\varphi(\vec{A})$ by some other, asymmetric function, $\varphi(\vec{A}) \to \tilde{\varphi}(\vec{A})$. Unfortunately, in the general case there is no obvious prescription how to achieve this. One possible way is to apply Gaussian approximation to $\varphi(\vec{A})$ in the vicinity of one of the equivalent equilibrium solutions, i.e., $\vec{A} = \vec{A}_0$. Under certain conditions this step is legitimate, as discussed below. On the other hand, we may argue that if other regions of $\mathcal{D}_{\mathcal{A}}$ are also characterized by large values of $\varphi(\vec{A})$ (i.e., points in the vicinity of other degenerate minima of $\mathcal{F}_z(\vec{A})$), they are separated by very large energy barriers, and hence practically inaccessible for the system characterized by $\vec{A} = \vec{A}_0$.

This reminds us, that the problem of degenerate states, characterized by the same values of (generalized) thermodynamic potentials and leading to vanishing averages, is present not only in non-MF statistical physics, but also within the MF treatment, where it is only somehow 'swept under the carpet'. Obviously, this is not only the property of the microscopic MF models, it is also present in the Landau theory.

12.1.4 Dual nature of fluctuations

Having both $\hat{\rho}_z(\vec{A})$ and $\varphi(\vec{A})$ present in our formalism, we may consider two distinct kinds of fluctuations. Apart from the above mentioned 'classical', there appear also 'quantum' fluctuations. The magnitude of the former is proportional to second moments of $\tilde{\varphi}(\vec{A})$, i.e., to $\overline{A_s A_t} - \overline{A_s} \cdot \overline{A_t}$. For a given $\vec{A}$, the magnitude of the latter is proportional to $\mathrm{Tr}[(\hat{A}_s - A_s)^2 \hat{\rho}_z(\vec{A})]$. It turns out, than under certain assumptions about the spatial dependence of the mean-fields, this dual picture becomes simpler in the thermodynamic ($\Lambda \to \infty$) limit.

12.1.5 Classical fluctuations: some definitions and notation

To proceed further, we have to analyze a dependence of $\mathcal{F}_z(\vec{A})$ on Λ, in particular in the $\Lambda \to \infty$ limit. To this end, we have to assume something about the Λ-dependence of a number of mean-fields M. We have to distinguish two situations; first, $M(\Lambda) = \mathrm{const}$, and second, when $M(\Lambda)$ depends non-trivially on Λ. In the latter case we do not have one function $\mathcal{F}_z(\vec{A})$, but rather a family of such functions, $\mathcal{F}_z^{(\Lambda)}(A_1, \ldots, A_{M(\Lambda)})$, labeled by Λ.

Within a general notation used up to this point, for $s \neq s'$, A_s and $A_{s'}$ may correspond both to different physical quantities, or to different component of some vector quantity, or finally to values of the same scalar quantity (or component of the vector) defined on different lattice sites. For the purpose of the present discussion it is necessary to label the mean-fields in a more precise manner, which allows to take their spatial dependence into account explicitly. Without the loss of generality we consider two-dimensional rectangular lattice.

Let our lattice be a rectangle of $\Lambda = \Lambda_x \Lambda_y$ sites. We divide it into $N_B = N_{Bx} N_{By}$ blocks of $n_B = n_{Bx} n_{By}$ sites each, hence $\Lambda = N_B n_B$. By l_x (l_y) we denote the lattice constant, whereas by L_x (L_y) the length of a side of each block along the x (y) direction, respectively. Therefore we have $L_x = n_{Bx} l_x$ and similarly for y-direction. It is now convenient to label each lattice site by a double vector index $\vec{I} = (\vec{R}, \vec{r})$, where $\vec{R}$ is a 'super-lattice' vector of a given block, and $\vec{r}$ labels sites within each such block. Therefore, $\vec{R} + \vec{r}$ is a position vector of a given lattice site labeled by $\vec{I}$. We also have $\vec{r} \equiv k_x l_x \hat{x} + k_y l_y \hat{y}$, $\vec{R} \equiv K_x L_x \hat{x} + K_y L_y \hat{y}$, where $0 \leq k_x \leq n_{Bx}$ and $0 \leq K_x \leq N_{Bx}$ and analogously for y direction (by $\hat{x}$ ($\hat{y}$) we denote the respective unit vectors). Note, that we usually assume periodic boundary conditions. Now we may write the mean-field index as $s = (q, \vec{I}) = (q, (\vec{R}, \vec{r}))$ where $q = 1, \ldots, N_Q$ labels different physical quantities (or components of vector quantities) and $\vec{I}$ is a lattice - site index introduced above.

12.1.6 Constant number of mean-fields

First, let us consider the case $M(\Lambda) = \text{const}$. Additionally, we assume that our system is translationally invariant with respect to the 'super-lattice' vectors $\vec{R}$, i.e., the corresponding sites in different blocks (i.e., having the same $\vec{r}$) are equivalent. This implies, that the number of independent MF variables M is bounded, $M \leq N_Q n_B$ (some mean-fields may be characterized by higher spatial symmetry then others). In other words, as a consequence of the constant value of M there exists an exact long-range order in our system. At least those sites are perfectly correlated, which are connected by the super-lattice vector, i.e., for any K_x and K_y we have

$$a(q, \vec{R}, \vec{r}) = a(q, \vec{R} + K_x L_x \hat{x} + K_y L_y \hat{y}, \vec{r}) \equiv a(q, \vec{r}). \tag{12.256}$$

In above, we used 'a' instead of 'A' to stress, that $a(q, \vec{R}, \vec{r})$ quantities are local, i.e., defined on a particular site. Note that the condition (12.256) introduces the 'infrared cutoff' in our model, apart from the 'ultraviolet cutoff' naturally appearing in any lattice model. Moreover, it is convenient to define the following global mean-field variables as

$$A(q, \vec{r}) = \sum_{\vec{R}} a(q, \vec{R}, \vec{r}) = N_B a(q, \vec{r}) = \frac{\Lambda}{n_B} a(q, \vec{r}). \tag{12.257}$$

If we re-scale the lattice size by a factor γ, i.e., $\Lambda \to \gamma\Lambda$, we have indeed

$$a(q, \vec{R}, \vec{r}) \to a(q, \vec{R}, \vec{r}), \qquad A(q, \vec{r}) \to \gamma A(q, \vec{r}). \tag{12.258}$$

Mean-fields $A(q, \vec{r})$ given by (12.257), which are exactly extensive from definition, will be called *spatially periodic*.

Now let us fix value of each $A(q, \vec{r})$, and re-scale the lattice as in (12.258). For finite Λ, $\mathcal{F}_z(\vec{A})$ is usually not strictly extensive. For example, for free-fermion models, by increasing the volume we also change the periodic boundary conditions, hence the allowed values of the wave vector. Consequently also the value of $\mathcal{F}_z(\vec{A})$ per site is altered. However, usually this departure from extensive behavior is expected to be small (presumably of order $1/\Lambda$) and vanish for $\Lambda \to \infty$. Thus we may write

$$\mathcal{F}_z(\vec{A}; \gamma\Lambda) \approx \gamma \mathcal{F}_z(\vec{A}; \Lambda), \qquad \lim_{\Lambda\to\infty} \left(\frac{\mathcal{F}_z(\vec{A}; \gamma\Lambda)}{\mathcal{F}_z(\vec{A}; \Lambda)} \right) = \gamma \tag{12.259}$$

with the equality in the thermodynamic limit. As a consequence, in particular we expect the same behavior for $\vec{A} = \vec{A}_0$, i.e., for the equilibrium values of mean-fields. Therefore, in the $\Lambda \to \infty$ limit, thermodynamic potentials (i.e., grand potential Ω, as defined in the next Section) should become extensive quantities.

Obviously, spatial dependence of the equilibrium values of mean-fields may be characterized by a higher translational (or rotational) symmetry than we initially assumed, with the homogeneous solution as a limiting case. Nonetheless, the symmetry-breaking patterns within each block may be also quite complicated, as those encountered for stripe phases (cf. e.g. Refs. [73, 75, 158]) or valence-bond solid phase [78]. Note, that not only in the latter case, but also even for a simple Neel antiferromagnetic state, the order parameter changes considerably on the atomic length scale.

If the condition (12.259) is obeyed, from (12.257) and (12.258) it follows that for the equilibrium (subscript 0) values of the extensive (global) mean-fields we have

$$\lim_{\Lambda\to\infty} \Lambda^{-1} A(q, \vec{r})_0 = \lim_{\Lambda\to\infty} \frac{1}{n_B} a(q, \vec{r})_0 > 0. \tag{12.260}$$

Now we are in a position to analyze classical fluctuations of those mean-fields, which are characterized by the specific periodic spatial dependence defined by (12.256). First we have to determine the actual equilibrium values of mean fields, $\vec{A} = \vec{A}_0$. We define $f_z(\vec{A}) \equiv \mathcal{F}_z(\vec{A})/\Lambda$. For large values of Λ, in most cases $f_z(\vec{A})$ should be practically Λ-independent (cf. Eq. (12.259)). Then, in order to obtain a Gaussian approximation for $\varphi(\vec{A}) \to \tilde{\varphi}(\vec{A}) \approx \varphi(\vec{A})_G$, we may expand $f_z(\vec{A})$ up to the second order in the components of $\vec{A} - \vec{A}_0$ around $\vec{A}_0$. This yields

$$\varphi_G(\vec{A}) = \mathcal{N}_G^{-1} \exp\Big(- \beta\Lambda\big\{ f_z(\vec{A}_0) + \frac{1}{2}\big(\vec{A} - \vec{A}_0\big)^T \mathcal{M}(\vec{A}_0)\big(\vec{A} - \vec{A}_0\big)\big\}\Big). \tag{12.261}$$

In above, $\mathcal{N}_G$ is the normalization constant, given by

$$\mathcal{N}_G = \int_{\mathcal{D}_\mathcal{A}} e^{-\beta\Lambda\left\{ f_z(\vec{A}_0) + \frac{1}{2}\left(\vec{A}-\vec{A}_0\right)^T \mathcal{M}(\vec{A}_0)\left(\vec{A}-\vec{A}_0\right)\right\}} d\vec{A}. \tag{12.262}$$

By $\mathcal{M}(\vec{A})$ we denote matrix of the second derivatives of $f_z(\vec{A})$, i.e.,

$$\Big(\mathcal{M}(\vec{A})\Big)_{tw} = \frac{\partial^2 f_z(\vec{A})}{\partial A_t \partial A_w}. \tag{12.263}$$

Obviously, for the true equilibrium, $\mathcal{M}(\vec{A}_0)$ should be positive-definite. Then, for large Λ, it is legitimate to extend the integration on the r.h.s. of (12.262) behind $\mathcal{D}_\mathcal{A}$, including unphysical values of $\vec{A}$. This is allowed as long as such $\vec{A}$ have a vanishing weight, and then the multiple integral (12.262) can be easily computed. In the $\Lambda \to \infty$ limit, regardless the value of temperature, $\varphi_G(\vec{A})$ approaches delta distribution centered at $\vec{A}_0$. Consequently, the second (and higher) moments of $\varphi_G(\vec{A})$ vanish for large Λ, and the 'classical' fluctuations of order parameters could be entirely neglected. This is also the necessary condition for the validity of the equilibrium description, constructed in Subsection 4.6. In other words, apart from the close vicinity of each of the phase transition points, those $\vec{A}$ which may occur with non-negligible probability, are practically indistinguishable from $\vec{A}_0$. Other minima, either local or global, are assumed to be inaccessible due to very large energy barriers separating them. On the other hand, near the transition points ($A_{s0} \approx 0$), Gaussian approximation cannot be valid, and we encounter all the problems appearing in the original Landau approach.

Let us point out once more, that the above conclusions are valid, provided that the condition (12.256) is fulfilled. This, however, is equivalent to the introduction of the infrared cutoff in our model 'by hand'.

12.1.7 Spatial dependence of mean-fields: general case

Now let us consider a general situation, i.e., when the value of (at least) one of the mean-fields on each site is allowed to be an independent variable. Consequently, we have $M \sim \Lambda$, and some of the conclusions drawn in the previous Sub-subsection for the case of spatially periodic mean-field configurations may be no longer valid now. In the present situation, (12.256) does not hold and fluctuations of arbitrary long-wavelength may be present in the system. Consequently, there may be either no true long-range order, or it may be essentially weakened.

When considering a particular equilibrium MF configuration $\vec{A} = \vec{A}_0$, we have to estimate its weight relative to the non-equilibrium ($\vec{A} \neq \vec{A}_0$) configurations. In other words, for a given MF model, we have to investigate the influence of classical fluctuations of the mean-fields on the equilibrium situation. This may depend strongly on e.g. the lattice geometry, value of external fields or temperature, but obviously also on the particular MF Hamiltonian. Therefore,

below we can provide only some qualitative analysis.[61] Still, it seems that there are two main possibilities.

First, if only those MF configurations have a non-negligible weight, which are practically indistinguishable from the equilibrium configuration. Let us introduce some vector norm $|\vec{A}|$ (its explicit form is not required here) in order to quantify the notion of distance between different mean-field configurations. Let ε be such a small number, than we consider two spatial mean-field configurations as practically indistinguishable[62] if

$$|\vec{A} - \vec{A}_0| < \varepsilon \max(|\vec{A}_0|, |\vec{A}|) \tag{12.264}$$

Then, for the situation considered here, the following condition must be satisfied

$$|\vec{A} - \vec{A}_0| \geq \varepsilon \cdot \max(|\vec{A}_0|, |\vec{A}|) \quad \Longrightarrow \quad \mathcal{F}_z(\vec{A}) \gg \mathcal{F}_z(\vec{A}_0) \Leftrightarrow \tilde{\varphi}(\vec{A}) \ll \tilde{\varphi}(\vec{A}_0). \tag{12.265}$$

This means that $\tilde{\varphi}(\vec{A})$ have to be strongly peaked at $\vec{A} = \vec{A}_0$. However, one may argue, that the above condition alone is not sufficient for the equilibrium situation to be well-defined, i.e., not destroyed by the classical fluctuations. It may seem more natural to require much stronger condition, namely, that

$$\int_{\mathcal{D}_{\mathcal{A}} \setminus \mathcal{K}(\vec{A}_0, R(\vec{A}_0))} \tilde{\varphi}(\vec{A}) d\vec{A} \quad \ll \quad \int_{\mathcal{K}(\vec{A}_0, R(\vec{A}_0))} \tilde{\varphi}(\vec{A}) d\vec{A} \sim 1, \tag{12.266}$$

with $\mathcal{K}(\vec{A}_0, R(\vec{A}_0)) \equiv \{\vec{A} : |\vec{A} - \vec{A}_0| = R(\vec{A}_0)$, where $R(\vec{A}_0)$ is defined as follows: the MF configurations are macroscopically indistinguishable if $|\vec{A} - \vec{A}_0| < R(\vec{A}_0)$. Only in such case we deal with a practically single, well defined value of $\vec{A}$. This is a necessary condition for a validity of the MF description.

Apparently, (12.266) have no precise meaning unless we specify the $\tilde{\varphi}(\vec{A})$ function. We will not comment here on the problem of construction of $\tilde{\varphi}(\vec{A})$ from the original probability distribution $\varphi(\vec{A})$ (12.254). In the previous sub-subsection, for $M = \text{const}$, $\tilde{\varphi}(\vec{A})$ has been constructed in the form of a Gaussian approximation, $\tilde{\varphi}(\vec{A}) = \varphi_G(\vec{A})$ (12.261).

Obviously, the condition (12.266) cannot be satisfied if we replace $\tilde{\varphi}(\vec{A})$ by $\varphi(\vec{A})$. Fortunately, even if only a weaker condition (12.265) is satisfied, we may once again invoke the argument, that values of $\vec{A}$, in the vicinity of other degenerate minima of $\mathcal{F}_z(\vec{A})$) are practically inaccessible due to large energy barriers. Then, $\varphi(\vec{A})$ may be replaced by $\tilde{\varphi}(\vec{A})$ of a Gaussian form, even if the reasoning leading to Eq. (12.261) can no longer be applied now.

Summarizing, we may hope, that for any possible reasonable form $\tilde{\varphi}(\vec{A})$ (not necessarily a Gaussian), obeying (12.265) the situation is qualitatively similar to that with $M = \text{const}$ and spatially periodic mean-fields, i.e., the equilibrium situation is stable against classical fluctuations. Note, that from Eq. (12.250) it follows, that for a given model, the above behavior is more likely to take place at *lower* temperatures.

One may argue that we probably should average the MF-configurations over many sites within some 'coarse-graining' procedure (i.e., by introducing block variables in analogy to the methods of Renormalization Group theory). However, as already mentioned, within the present formalism we are not allowed to average over classical field configurations. The averaging procedure is performed only over the quantum, microscopic degrees of freedom, labeled by eigenstates $|i(\vec{A})\rangle$ of the MF Hamiltonian.

If (12.266) or the weaker condition (12.265) is not satisfied, the mean-field theory breaks down. In particular, these conditions are expected to be violated in the vicinity of the phase

[61] Some insight may be gained from the analysis of the MF Ising model, cf. Sec. 12.3.

[62] Obviously, the phrase 'practically indistinguishable' is imprecise. In practice, value of ε depends e.g. on the details of the experimental situation we try to describe using our formalism.

transition points. Moreover, within the present formulation, there exist yet another natural limitation of the MF approach, imposing additional restrictions for the allowed MF configurations. This is analyzed in the next Sub-subsection.

12.1.8 Quantum fluctuations

Presence of 'quantum' probability distribution given by $\hat{\rho}_z(\vec{A})$ (4.42) allows us to look at the internal consistency of the MF approach from yet another perspective. Namely, it seems that a natural necessary condition for the validity of such semi-classical description is that quantum fluctuations of observables $\hat{A}_1, \ldots, \hat{A}_M$ are negligible. More precisely, for any $\hat{A}_s$, we should require that the quantum dispersion $\sigma(\hat{A}_s)$, i.e

$$\sigma(\hat{A}_s) = \sqrt{\mathrm{var}(\hat{A}_s)} = \sqrt{\mathrm{Tr}[(\hat{A}_s - A_s)^2 \hat{\rho}_z(\vec{A})]}, \tag{12.267}$$

is small as compared to the expectation value of $\hat{A}_s$. In other words, we are allowed to treat mean-field variables as semi-classical quantities only for those points $\vec{A} \in \mathcal{D}_{\mathcal{A}}$, for which the following condition

$$\sigma_r^2(\hat{A}_s) \equiv \frac{\mathrm{var}(\hat{A}_s)}{(A_s)^2} \ll 1, \tag{12.268}$$

is satisfied (with $A_s \neq 0$). Therefore, (12.268) should be fulfilled at least in the vicinity of the equilibrium solution, $\vec{A} = \vec{A}_0$.

Similarly to the situation encountered in the case of 'classical' fluctuations, the relative role of quantum fluctuations, i.e., the behavior of the l.h.s. of (12.268) crucially depends on the spatial configuration of mean-fields.

For $M = \mathrm{const}$, $\sigma_r^2(\hat{A}_s)$ should vanish in the $\Lambda \to \infty$ limit as $\sim 1/\Lambda$. This is expected to hold for the most lattice mean-field models, under quite general assumptions, e.g finite range of interactions. Note, that for a given value of $\vec{A}$, this may be verified in a way essentially identical to that for the non-MF models. Obviously, (12.268) can never be satisfied for $A_s = 0$, as well as for finite, but small values of A_s, $A_s \approx 0$ if $\Lambda < \infty$. However, for nonzero, albeit small A_s we always expect $\sigma_r^2(\hat{A}_s)$ to eventually vanish in the thermodynamic $\Lambda \to \infty$ limit. Therefore we may expect that the 'macroscopic' mean-fields obeying (12.258) are not destroyed by quantum fluctuations, and for such mean-fields condition (12.268) is satisfied.

On the other hand, for the solutions characterized by an arbitrary spatial dependence of mean-fields, $\sigma_r^2(\hat{A}_s)$ may be of the order of unity or larger, regardless the value of Λ, and even in the $\Lambda \to \infty$ limit. An example is provided by MF Ising model, cf. Supplement C, Sub-subsection 12.3.6. Nonetheless, we may expect, that arbitrary MF configuration $\vec{A}$ is protected against quantum fluctuations as long as it does not depart to much from some spatially-periodic mean-field configuration $\vec{A}^{(p)}$ obeying (12.264). With the increasing $|\vec{A} - \vec{A}^{(p)}|$, quantum fluctuations of at least one mean-field may become pronounced. Eventually such fluctuation may be of the same order of magnitude as $|A_s - A_s^{(p)}|/|\vec{A}_s^{(p)}|$.

Finally, in the connection with (12.268) we may invoke discussion provided in Chapter XII of Ref. [148], where the conditions, under which physical quantity x may be assumed to behave classically, are analyzed. Quantity x is assumed to have vanishing average value, $\bar{x} = 0$, and this averaging procedure seems to correspond to our classical average, defined with the help of $\varphi(\vec{A})$. Thus, we may identify x with each of the quantities $A_s - A_{s0}$. Using arguments based on the time-energy uncertainty principle, Landau and Lisfshitz arrive at the following condition (Eq. (110.2) of [148])

$$T \gg \hbar/\tau, \quad \tau \gg \hbar/T, \tag{12.269}$$

with τ being a rate of change of x, i.e., $\dot{x} \sim x/\tau$. Clearly, if τ is finite, the temperature cannot be too low. However, here due to conditions (4.18), we rather have $\tau = \infty$. Therefore, it seems

that arbitrary low temperature can be consistently treated within our approach. However, in the present case, the criterion for the mean-field approach validity should be based not on the condition (12.269), but rather on (12.267) and (12.268). Please note, that the condition (12.268) is expected to be satisfied better at lower temperatures. This is because in the $T \to 0$ limit, quantum fluctuations usually decrease, whereas the magnitude of each of the order parameters in equilibrium increases.

12.1.9 Fluctuations: summary and final remarks

Summarizing, for a given MF model, the values of the model parameters and thermodynamic variables, the range of applicability of the MF formalism (i.e., its internal consistency) is determined by two factors. First, a particular MF configuration $\vec{A}$ may violate condition (12.265). In such a situation the equilibrium is unstable against 'classical' fluctuations. Second, for some $\vec{A}$, 'quantum' fluctuations may be so pronounced, that they invalidate the assumption of all mean-fields being well-defined, static quantities. In other words, such MF configurations $\vec{A}$ are destroyed by the quantum fluctuations.

However, quantum fluctuations should be negligible in the thermodynamic limit, if only each of the mean-fields A_s obeys the condition (12.257). Then, each mean-field is well-defined for large Λ and for arbitrary temperature $T < T_c^{(s)}$, except a close vicinity of the corresponding phase transition point $T = T_c^{(s)}$. Also, we may expect that large values of the (macroscopic) order parameters imply small influence of quantum fluctuations. Again, this may be observed on the example of the MF Ising model of Sec. 12.3 and the spatial configurations of magnetization which are close to the uniform, saturated configuration of order parameter.

Analysis of both classical and quantum fluctuations leads us to the following general conclusion. Mean-field, semi-classical description used here works better at low temperatures. With the increasing temperature, MF description is invalidated by the fluctuations of both classical and quantum nature. This may be surprising at the first glance, because we usually associate classical behavior with higher temperatures, and expect genuine quantum effects to appear at lower temperatures. Nonetheless, the above conclusion is consistent with the previously invoked remark made in Ref. [158], attributing the validity of MF description to nonzero value(s) of order parameter(s). It is also consistent with the well-known tendency of the MF models to overestimate the range of ordered phases (and hence the values of the corresponding critical temperatures), which, in turn, is commonly attributed to the negligence of the thermodynamic fluctuations.

12.2 Supplement B: Zero temperature limit of mean-field approach

12.2.1 Introductory remarks

Mean-field models are frequently solved at $T = 0$, without any reference to a finite-temperature situation. Such analysis is usually based on the variational principle of quantum mechanics, cf. e.g. [66, 69, 70, 71, 79, 85, 137]. Obviously, exactly at $T = 0$ we cannot use the MaxEnt principle in the present formulation, i.e., based on the maximization of $\mathcal{S}_\lambda$ (4.14). However, validity of this approach is restored at any $T > 0$ ($\beta < \infty$). On the other hand, we need some elements of MF quantum-mechanical description of pure states at $T = 0$ in order to construct a finite-temperature mean-field approach, e.g. to justify statements concerning time evolution of quantum states, which have been made in Subsection 4.3. Thus, for the sake of completeness, below we provide a Reader with the particular formulation of the MF theory for $T = 0$.

However, we also point out deficiencies of any zero - temperature MF approach. We argue, that due to both technical and conceptual difficulties that are encountered, it is preferable to study $T = 0$ situation as the $T \to 0$ limit of finite-temperature case, i.e., using the MaxEnt-based formalism of the previous Subsections. This point of view is also supported by the observation, that for the most of the mean-field models of interest, the equilibrium values of the MF variables depend very weakly on the temperature near $T = 0$. From the point of view of numerical analysis, this is a convenient feature. On the other hand, this fact may indicate limited applicability of the MF methods for the description of quantum phase transitions.

12.2.2 Assumptions

Similarly to the $T > 0$ case, we want to base our description of the system entirely on the *a priori* given MF Hamiltonian $\hat{H}(\vec{A})$. Thus, we argue the self-consistency requirements as expressed by Eqs. (4.8) for the appropriate form of $\hat{\rho}(\vec{A})$ should be fulfilled. Also, we assume that the state of the system can be described by a state vector, i.e., it is a pure quantum state. Then the third law of thermodynamics is automatically satisfied, as the von Neumann entropy of the corresponding density operator is equal zero. Moreover, again, we consider only time-independent situation. Construction of the non-trivially time-dependent MF formalism requires separate analysis and is postponed for further studies. Consequently, any pure MF quantum state we considered is an eigenstate of the MF Hamiltonian, or a coherent superposition of degenerate eigenstates, characterized with the same value of energy. This ensures that the time evolution of such quantum state is trivial.[63]

It seems natural to base the whole approach on the variational principle of quantum mechanics. It is well known, that for a non-MF Hamiltonian $\hat{H}_e$, the time-independent Schrödinger equation, $\hat{H}_e|\Psi\rangle = E|\Psi\rangle$ may be derived from the variational principle [223], namely by varying the following functional

$$\Phi_e(\Psi) = \langle\Psi|\hat{H}_e - E|\Psi\rangle \tag{12.271}$$

with respect to $\langle\Psi|$ (variation with respect to $|\Psi\rangle$ yield a conjugate equation). The normalization constraint $\langle\Psi|\Psi\rangle = 1$ is imposed by means of the Lagrange multiplier E, which becomes equal to one of the eigenvalues of $\hat{H}_e$.

At this point, we *assume*, that the same variational principle (after some necessary modifications, see below) may be also used in the context of the MF models, providing a convenient way to describe time-independent, situation at $T = 0$.

12.2.3 Incomplete approach

In order to apply the variational principle based on (12.271) in the present situation, it seems natural to introduce the following functional

$$\Phi(\Psi, \vec{A}) = \langle\Psi|\hat{H}(\vec{A}) - \mu\hat{N}|\Psi\rangle - E\langle\Psi|\Psi\rangle. \tag{12.272}$$

To fix the average particle number $\langle\hat{N}\rangle$ at the desired value, we have added the $(-\mu\hat{N})$ term to $\hat{H}(\vec{A})$. The presence of this term is usually unavoidable for the mean-field models, e.g. if

[63] In Subsection 4.3 we have assumed that time evolution of quantum states is given by Eq. (4.20). Exactly as in the non-MF case, Eq. (4.20) follows from the time-dependent MF Schrödinger equation,

$$(\hat{H}(\vec{A}) - \mu\hat{N})|\Psi(t)\rangle \equiv \hat{K}(\vec{A}, \vec{\lambda})|\Psi\rangle = i\partial_t|\Psi(t)\rangle, \tag{12.270}$$

provided that $\hat{K}(\vec{A}, \vec{\lambda})$ does not depend explicitly on time, and that mean-fields and Lagrange multipliers are time-independent. If the latter condition is not fulfilled, it remains unclear for us, if the time-dependent MF Schrödinger equation supplemented only with the conditions $A_s(t) = \langle\Psi(t)|\hat{A}_s|\Psi(t)\rangle$ is sufficient to describe consistently the time evolution of the system.

$[\hat{H}(\vec{A}), \hat{N}] \neq 0$, or when $\langle \hat{N} \rangle_z \equiv A_1$ is one of the mean-fields. If $A_1, \ldots A_M$ were not treated as variational parameters, and their implicit dependence on $|\Psi\rangle$ was neglected for a moment, (12.272) may be varied with respect to $\langle \Psi|$.[64] This yields the following time-independent Schrödinger equation

$$(\hat{H}(\vec{A}) - \mu \hat{N})|\Psi\rangle \equiv \hat{K}(\vec{A})|\Psi\rangle = E|\Psi\rangle. \tag{12.273}$$

By using Eq. (12.273), the ground state $|\Psi\rangle = |\Psi_0\rangle \equiv |\tilde{1}\rangle$ and corresponding ground state energy $\tilde{E}_1$ may be easily found. Optimal values of the mean-fields $A_1, \ldots A_M$ are then given by $A_{t(sc)} \equiv \langle \Psi_0 | \hat{A}_t | \Psi_0 \rangle$. This is in a full analogy with the finite-temperature case, where the similar treatment led to the density operator of the apparently canonical form (4.6) with $\hat{H}_e$ replaced by $\hat{H}(\vec{A})$. Again, at $T = 0$ such an approach seems to be incomplete, because the implicit dependence of $\vec{A}$ on $|\Psi\rangle$ has been neglected. In other words, in contrast to the case of a non-MF quantum mechanics, Schrödinger equation of the form (12.273) does not follow from the variational principle.

12.2.4 The method of Lagrange multipliers

Consequently, we proceed still in a direct analogy with the analysis carried out in Subsections 4.3 and 4.4. We augment $\hat{H}(\vec{A})$ with the self-consistency preserving constraints, which allow to treat $A_1, \ldots A_s$ and $\langle \Psi|$ as independent variables. This step, in turn, yields again the MF Hamiltonian $\hat{H}_\lambda$ given by (4.15), which replaces $\hat{H}(\vec{A})$. Explicitly, the functional to be minimized with respect to both $\langle \Psi|$ and $A_1, \ldots A_s$ reads now

$$\Phi_\lambda(\Psi, \vec{A}, \vec{\lambda}) = \langle \Psi | \hat{H}(\vec{A}) - \sum_s \lambda_s (\hat{A}_s - A_s) - \mu \hat{N} | \Psi \rangle - E \langle \Psi | \Psi \rangle. \tag{12.274}$$

Varying (12.274) with respect to $\langle \Psi|$ we obtain time-independent Schrödinger equation

$$\hat{K}_\lambda(\vec{A}, \vec{\lambda})|\Psi\rangle = E|\Psi\rangle, \tag{12.275}$$

with

$$\hat{K}_\lambda(\vec{A}) \equiv \hat{H}(\vec{A}) - \sum_s \lambda_s (\hat{A}_s - A_s) - \mu \hat{N}. \tag{12.276}$$

We see, that indeed $\hat{K}_\lambda(\vec{A}, \vec{\lambda})$ rather than $\hat{H}(\vec{A})$ or $\hat{H}(\vec{A}) - \mu \hat{N}$ should be identified with the proper MF Hamiltonian. Taking further the derivatives with respect to A_s and λ_s, we obtain

$$\left\langle \Psi \left| \left(\frac{\partial \hat{H}(\vec{A})}{\partial A_s} + \lambda_s \right) \right| \Psi \right\rangle = 0, \tag{12.277}$$

i.e., the $T = 0$ limit of Eqs. (4.30), as well as

$$\langle \Psi | \hat{A}_s | \Psi \rangle - A_s \equiv \langle \hat{A}_s \rangle_\Psi - A_s = 0, \tag{12.278}$$

corresponding to Eqs. (4.31). If additional variational parameters $b_1, \ldots, b_P$ of a non-MF character are present, we minimize (12.274) also with respect to each b_l. This yields

$$\left\langle \Psi \left| \left(\frac{\partial \hat{H}(\vec{A}, \vec{b})}{\partial b_l} \right) \right| \Psi \right\rangle = 0, \tag{12.279}$$

corresponding to Eqs. (4.27). Solution of the initial problem is determined by Eqs. (12.275)-(12.279), from which we obtain $\vec{A} = \vec{A}_0$, $\vec{\lambda} = \vec{\lambda}_0$, $\vec{b} = \vec{b}_0$ and the ground state $|\Psi\rangle = |\Psi_0\rangle \equiv |1\rangle$. Still, below we provide the Reader with additional discussion and further remarks that are in place at this point.

[64] By a derivative of $\Phi_\lambda(|\Psi\rangle, \vec{A}, \vec{\lambda})$ with respect to $\langle \Psi|$ we understand, as usual, a vector of partial derivatives of Φ_λ with respect to c_l^*, where c_l are the expansion coefficients of $|\Psi\rangle$, i.e., $|\Psi\rangle = \sum_l c_i |l\rangle$ in the basis of $|l\rangle_1^D$.

12.2.5 $\vec{A}, \vec{\lambda}$ - independent eigenstates of $\hat{H}_\lambda(\vec{A})$

By solving only Eq. (12.275), we obtain eigenvalues $E_i(\vec{A}, \vec{\lambda})$ and eigenvectors $\{|i(\vec{A}, \vec{\lambda})\rangle\}_{i=1}^{D_H}$ of $\hat{K}_\lambda(\vec{A})$. This remains in a complete analogy to the $T > 0$ case, where Eqs. (4.23) and (4.24) have been used to obtain explicit $\vec{A}$ and $\vec{\lambda}$ - dependence of the density operator $\hat{\rho}_\lambda(\vec{A}, \vec{\lambda})$ (4.29). However, here we have to distinguish between the two situations. In general, the eigenbasis of $\hat{H}_\lambda(\vec{A})$ may be non-trivially $\vec{A}$ or $\vec{\lambda}$-dependent.[65] Then, at least in principle, we should be able to solve Eqs. (12.278) for each $\vec{A}$ and i, with $|\Psi\rangle = |i(\vec{A}, \vec{\lambda})\rangle$, and we obtain $\lambda = \vec{\lambda}_i(\vec{A})$ (note that Eqs. (12.278) may be regarded as a $T \to 0$ limit of Eqs. (4.70)). Consequently, we should be able to obtain D_H functions $E_i^{(z)}(\vec{A}) \equiv E_i(\vec{A}, \vec{\lambda}_i(\vec{A}))$. We are mainly interested in the $\vec{A}$-dependence of the ground-state energy, $E_1^{(z)}(\vec{A})$, which seems to correspond to the 'classical energy landscape' of Ref. [157].

However, eigenstates of $\hat{H}_\lambda(\vec{A})$ may be also $\vec{A}, \vec{\lambda}$ - independent, e.g. this is the case if we have $[\hat{H}_\lambda(\vec{A}), \hat{A}_s] = 0$ for some s. In such case, $\vec{\lambda}$ does not appear on the l.h.s. site of (12.278), and therefore the latter equations cannot be solved for $\vec{\lambda}$. As a consequence, we cannot apply the formal results of Subsection 4.8, and the existence of universality classes and equivalence classes of MF Hamiltonians is obscured. Also, only discrete values of the mean fields, equal to $\langle i|\hat{A}_s|i\rangle$, are permitted and the self-consistency conditions (4.8) are fulfilled automatically. A discrete character of the mean fields is, however, unnatural.

Please note, that the calculus of variations and method of Lagrange multipliers are applicable, and Eqs. (12.275)-(12.278) remain valid in any case, regardless the detailed form of $\vec{A}, \vec{\lambda}$ - dependence of the eigenvectors of $\hat{H}_\lambda$. This is true as long as $\Phi_\lambda(\Psi, \vec{A}, \vec{\lambda})$ is differentiable with respect to all of its variables.

12.2.6 Non-analytical minima

Eqn. (12.277) provides a part of necessary conditions for the existence of the minimum of $\Phi_\lambda(\Psi, \vec{A}, \vec{\lambda})$ (12.274) prior to the self-consistency constraints, *provided* that Φ_λ is differentiable in the neighborhood of such minimum, i.e that the minimum lies in the interior of the domain $\mathcal{D}_\mathcal{A}$. This, however, is frequently not the case. True conditional minimum may lay on the boundary of $\mathcal{D}_\mathcal{A}$, and thus not correspond to the stationary point of $\Phi_\lambda(\Psi, \vec{A}, \vec{\lambda})$. In such case we are forced to examine carefully the boundary of $\mathcal{D}_\mathcal{A}$, but this may be difficult if we have to rely only on the numerical analysis.

This may be understood as follows. At $T = 0$, minimum of the ground-state energy as a function of $\vec{A}$ frequently corresponds to the minimal or maximal permitted value of at least one of the mean fields $A_1, A_2, \ldots, A_M$. Clearly, such minimum is located at the boundary of $\mathcal{D}_\mathcal{A}$. For example, this is the case for some MF models describing ferro- or antiferromagnetism, where the ground state may correspond to saturated magnetization with no spin-wave excitations. In particular, this situation is present in the MF Ising model, analyzed in Subsection 12.3, but possibly also in more complex MF models.

On the other hand, at $T > 0$ the entropic part of $\mathcal{F}$ favors thermally excited states. Consequently, at nonzero temperature, minimum of $\mathcal{F}$ is no longer located on the boundary of $\mathcal{D}_\mathcal{A}$ and its analytical character is restored. Then, Eq. (4.30), a finite-temperature counterpart of (12.277), together with the remaining equations, yield a correct solution. In the language of Landau theory, finite temperature restores the term $\sim \phi^4$, where ϕ denotes an order parameter in question. This is because for many microscopic MF models, the fourth-order term in the expansion of $\mathcal{F}$ in powers of the order parameter $A_s = \phi$ is proportional to T (cf. Eq. (12.301)

[65] By this we understand that the expansion coefficients of each state $|i(\vec{A}, \vec{\lambda})\rangle$ in a discrete coordinate basis (or in eigenbasis of any $\vec{A}, \vec{\lambda}$ - independent operator) depend in a nontrivial way on $\vec{A}, \vec{\lambda}$.

in Sec. 12.3)), i.e.,

$$\mathcal{F}(\phi) = \mathcal{F}_0 + A(T - T_c)\phi^2 + TB\phi^4 + O(\phi^6) \tag{12.280}$$

$A < 0$ and $B > 0$ are functions of other mean fields A_t, $t \neq s$. Above discussed feature of the $T = 0$ analysis is very inconvenient from the point of view of numerical analysis. Therefore, sometimes a finite-temperature MF formalism is introduced only in order to ensure better convergence of the numerical procedures, and to avoid the 'open shell' problem by replacing step function by the smooth Fermi distribution at $T > 0$, [73, 74, 75, 78].

12.2.7 Excited states

Another problem arises, when within the zero-temperature MF formalism, excited states are to be constructed. One may try to use again the variational principle based on (12.274). Then, the k-th excited state may be obtained by minimization of (12.274) within the subspace orthogonal to all previously found states of lower energy, i.e., $|\Psi_1\rangle, |\Psi_2\rangle, \ldots, |\Psi_{k-1}\rangle$, where $|\Psi_1\rangle$ denotes the ground state (no degeneracy is assumed for simplicity). It can be realized by adding to (12.274) $k-1$ additional constraints (again, to achieve this, the method of Lagrange multipliers may be employed)

$$\langle\Psi_k|\Psi_1\rangle = \langle\Psi_k|\Psi_2\rangle = \ldots \langle\Psi_k|\Psi_{k-1}\rangle = 0. \tag{12.281}$$

However, then, in general, each excited state is characterized by different optimal values of mean-fields $\vec{A} = \vec{A}_{0i}$ and Lagrange multipliers $\vec{\lambda} = \vec{\lambda}_{0i}$. On the other hand, quantum states characterized by different values of those semi-classical parameters are in fact eigenstates of *different* MF Hamiltonians, moreover, they should be interpreted as states from different Hilbert spaces, as discussed in Sub-subsection 4.11.2.

To avoid the above mentioned difficulty, one may use the ground-state values of the mean fields ($\vec{A} = \vec{A}_0$), Lagrange multipliers ($\vec{\lambda} = \vec{\lambda}_0$), and the variational parameters of non-MF character, ($\vec{b} = \vec{b}_0$), obtained from the minimization procedure given by Eqs. (12.275)-(12.278). In such case the eigenbasis may be defined as $|i\rangle \equiv |i(\vec{A}_0, \vec{\lambda}_0)\rangle$. This route corresponds to the $T \to 0$ limit of the $T > 0$ situation. Nonetheless, by doing so, we treat the ground and the excited states in an asymmetric manner; for $i = 2, \ldots, D_H$ we have $\langle i|\hat{A}_s|i\rangle \neq A_{s0} = \langle 1|\hat{A}_s|1\rangle$.

12.2.8 Summary: deficiencies of zero-temperature MF approach

Above we have mentioned several problems arising within the MF approach at $T = 0$. Note, however, that none of them is present within the corresponding finite-temperature, MaxEnt-based treatment. Namely, in the latter case, the equilibrium values of both $\vec{A}_0$ and $\vec{\lambda}_0$ are determined *collectively* by all quantum states present in the partition function $\mathcal{Z}_\lambda$ (4.29). This is true even at lowest (albeit finite) temperatures, when the overwhelming contribution to $\mathcal{Z}_\lambda$ comes from the ground state.

If $\hat{H}(\vec{A})$ depends on $\langle\hat{N}\rangle = A_1$ or if $[\hat{H}(\vec{A}), \hat{N}] \neq 0$, the $(-\mu\hat{N})$ term has to be introduced. As a consequence, pure quantum states may depend on μ, similarity as in the $T > 0$ case. It is, however, more natural to introduce μ in the context of a finite-temperature statistical-mechanical description. Therefore, as long as analysis of the time evolution of mean-fields is not required, $T = 0$ formulation of the MF formalism has no advantages over the finite-temperature, MaxEnt-based treatment. The $T \to 0$ limit of the latter provides a consistent MF description of the equilibrium properties of the system at $T = 0$, and allows to avoid caveats of the purely zero-temperature approach. Note also, that the variational principle used here, i.e., minimization of (12.274) has been postulated in a direct analogy with the case of a non-MF description. However, in the case of the MF models, its deeper justification is provided mainly by the fact, that results of this approach coincide with those of the MaxEnt-based finite-temperature treatment in the $T \to 0$ limit.

12.3 Supplement C: Mean-field model of the spin system as an illustrative example for application of the MaxEnt-based variational approach

On the example of a simple MF treatment of the Heisenberg [190] and Ising [133, 167] models, we illustrate here the method developed in Part II.

Within the simplest Hartree approximation for the interaction term, both models lead to the same MF Hamiltonian. The latter describes an ensemble of independent spins, interacting only with an effective field. Therefore, this approach is essentially a kind of a molecular field approximation [134, 135]. Interestingly, the results obtained here are formally identical to those of the Bragg-Williams approximation (BWA) [152] for the Ising model. Consequently, our method provides alternative derivation of BWA.

Due to the simplicity of the MF model analyzed here, for most of the quantities of interest explicit analytical formulas become available. This gives us a better insight into the structure of the present formalism and reveals some of its general features.

12.3.1 Construction of mean-field Hamiltonian

Consider a quantum Heisenberg model, defined on D_S-dimensional cubic lattice with Λ sites by the following Hamiltonian

$$\hat{H}_H = \sum_{\langle ij \rangle} J_{ij} \hat{\vec{S}}_i \cdot \hat{\vec{S}}_j - \vec{B}_i \cdot \hat{\vec{S}}_i. \tag{12.282}$$

In above, J_{ij} is an exchange integral, $\vec{B}_i$ is an external magnetic field vector on the i-th lattice site, and $\hat{\vec{S}}_i$ is a vector of Pauli matrices, $\hat{\vec{S}}_i = (\hat{\sigma}_i^x, \hat{\sigma}_i^y, \hat{\sigma}_i^z)$. The constants $\hbar^2/4$ and $\hbar/2$ have been incorporated in the definitions of J_{ij} and $\vec{B}_i$, respectively. Next, we apply the Hartree-type MF decoupling (4.107) to (12.282). This yields

$$\hat{H}_H \to \hat{H} = \sum_{\langle ij \rangle} J_{ij} \big(\hat{\vec{S}}_i \cdot \langle \hat{\vec{S}}_j \rangle + \langle \hat{\vec{S}}_i \rangle \cdot \hat{\vec{S}}_j - c \langle \hat{\vec{S}}_i \rangle \cdot \langle \hat{\vec{S}}_j \rangle \big) - \vec{B}_i \cdot \hat{\vec{S}}_i. \tag{12.283}$$

To illustrate certain properties of our approach, we have introduced $c \in [0, 1]$; $c = 1$ corresponds to standard Hartree approximation (4.107). Next, we supplement $\hat{H}$ with appropriate constraint terms and obtain MF Hamiltonian $\hat{H}_\lambda$ (4.15) of the form

$$\hat{H}_\lambda = \hat{H} - \sum_i \vec{\lambda}_i \cdot (\vec{S}_i - \langle \vec{S}_i \rangle). \tag{12.284}$$

Lagrange multipliers $\vec{\lambda}_i$ have a natural interpretation of the local molecular magnetic fields. We take $J_{ij} = -J < 0$ for the nearest neighbors, and $J = 0$ otherwise. Also, $\vec{B}_i = B\hat{z}$, $\vec{\lambda}_i = \lambda \hat{z}$, and $\langle \hat{\vec{S}}_i \rangle = \vec{S}_i = m\hat{z}$, as we are interested in the simplest solution, which exhibits full translational symmetry of the underlying lattice, and for which only z- components of corresponding vector fields retain nonzero values. The sum over bonds is equal half of the sum over sites, $\sum_{\langle ij \rangle} = \frac{1}{2} \sum_{i,j(i)}$, therefore summation over $j(i)$ gives just $z = 2d$, the number of nearest neighbors of a given site. All those assumptions and simplifications yield

$$\hat{H}_\lambda = -\big(Jzm + B + \lambda \big) \sum_i \hat{\sigma}_i^z + \Lambda \big(\frac{1}{2} Jzcm^2 + \lambda m \big). \tag{12.285}$$

By such a brutal approximation we have reduced (12.282) to an effective Hamiltonian describing the sum of decoupled spins, interacting with the combination of an applied external field and

molecular fields. Obviously, the same result could be obtained without some intermediate steps, if instead of (12.282), we chose as a starting point the Ising Hamiltonian,

$$\hat{H}_I = \sum_{\langle ij \rangle} J_{ij} \hat{\sigma}_i^z \hat{\sigma}_j^z - \sum_i B_i \hat{\sigma}_i^z. \tag{12.286}$$

12.3.2 Free energy functional and equilibrium situation

It is a simple exercise to compute the generalized free energy potential (4.33) for the Hamiltonian (12.285). We obtain

$$\mathcal{F}(m, \lambda; \beta, B) = \Lambda\Big\{ \frac{Jz}{2} c m^2 + \lambda m - \beta^{-1} \ln[2 \cosh\big(\beta(Jzm + B + \lambda)\big)] \Big\}. \tag{12.287}$$

Using (12.287), we obtain equilibrium values of the magnetization ($m = m_0$), and molecular field ($\lambda = \lambda_0$) by solving Eqs. (4.34), which in the present case read

$$\frac{\partial \mathcal{F}(m, \lambda)}{\partial \lambda} = \Lambda\big\{ m - \tanh\big(\beta(Jzm + B + \lambda)\big) \big\} = 0, \tag{12.288}$$

$$\frac{\partial \mathcal{F}(m, \lambda)}{\partial m} = \Lambda\big\{ (Jzcm + \lambda) - Jz \tanh\big(\beta(Jzm + B + \lambda)\big) \big\} = 0. \tag{12.289}$$

From (12.288) and (12.289) it follows that

$$\lambda_0 = Jz(1 - c)m_0, \tag{12.290}$$

and, consequently, m_0 is a solution of the following equation

$$m_0 - \tanh\big(\beta(Jz(2 - c)m_0 + B)\big) = 0. \tag{12.291}$$

λ_0 can be also obtained using (4.30), i.e.,

$$\lambda_0 = -\Big\langle \frac{\partial \hat{H}}{\partial m} \Big\rangle_{m=m_0} = -Jz(c - 1)m_0, \tag{12.292}$$

which is in agreement with Eq. (12.290). Clearly, for $c = 1$, $\lambda_0 = 0$, as expected for HF type of MF models (cf. discussion in Subsection 4.9). The critical temperature (Curie temperature) T_c can be now evaluated in a standard manner [136], and for $B = 0$ we have

$$k_B T_c = (2 - c)Jz. \tag{12.293}$$

For $c = 0$, T_c is two times larger then a standard MF value (obtained for ($c = 1$)), as without subtraction of the $\sim m^2$ term we largely overestimate the magnetic interactions, which are already severely overestimated (as compared to an exact solution) by the mean-field approximation (12.283).

12.3.3 Non-equilibrium situation

Let us consider now a non-equilibrium situation. From the self-consistency equation (12.288), for $m \in (-1, 1)$, we obtain

$$\lambda(m) = \beta^{-1} \mathrm{artanh}(m) - Jzm - B. \tag{12.294}$$

Therefore, $\hat{H}_z(m) = \hat{H}_\lambda(m, \lambda(m))$ (cf. Eq. (4.42) of Sec. 4.8) reads now

$$\hat{H}_z(m) = -\frac{\mathrm{artanh}(m)}{\beta} \sum_i \hat{\sigma}_i^z + \Lambda\Big(\frac{Jz}{2}(c-2)m^2 - Bm + m\frac{\mathrm{artanh}(m)}{\beta}\Big). \tag{12.295}$$

From (12.295) we see, then when the requirements of self-consistency are fulfilled, the coupling of observables (in the present case $\hat{S}^z_{tot} = \sum_i \hat{\sigma}_i^z$, the z-component of the total spin) to the mean fields (m) apparently changes, as compared to the initial MF Hamiltonian

$$\hat{H}(m) = \hat{H}_\lambda(m, \lambda(m) = 0) = -\big(Jzm + B\big) \sum_i \hat{\sigma}_i^z + \frac{\Lambda}{2} Jzcm^2, \tag{12.296}$$

in which the self-consistency requirements are ignored. Note also, that although (12.295) is not defined for $m = \pm 1$, appropriately defined limit $\beta \to \infty$, $m \to \pm 1$ of the $\beta^{-1}\mathrm{artanh}(m)$ term exists, as we will be discussed in what follows.

Landau free energy potential (cf. Eq. (4.44)), $\mathcal{F}_z(m) = \mathcal{F}(m, \lambda(m))$ may be now easily constructed. After simple manipulations, making use of the following identities

$$\cosh\big(\mathrm{artanh}(m)\big) = \frac{1}{\sqrt{1-m^2}}, \quad \mathrm{artanh}(m) = \frac{1}{2}\ln\left(\frac{1+m}{1-m}\right), \tag{12.297}$$

and defining

$$U(m) \equiv \langle \hat{H}_\lambda \rangle = \langle \hat{H}_z \rangle = \langle \hat{H} \rangle = \Lambda\big(\frac{1}{2} Jz(c-2)m^2 - Bm\big), \tag{12.298}$$

$$S(m) = \frac{1+m}{2}\ln\left[\frac{1+m}{2}\right] + \frac{1-m}{2}\ln\left[\frac{1-m}{2}\right]. \tag{12.299}$$

we obtain

$$\mathcal{F}_z(m) = U(m) - \Lambda\beta^{-1} S(m). \tag{12.300}$$

Remarkably, the above formula is identical to that obtained within Bragg - Williams approximation (BWA) for the Ising model [184]. However, our result is valid for arbitrary value of m and Λ. Also, Stirling approximation has not been used explicitly[66] to derive (12.300). In other words, BWA respects the self-consistency somewhat 'by accident', and only in the $\Lambda \to \infty$ limit, whereas within our method the self-consistency requirements are fulfilled by construction.

Landau free energy potential (12.300) has the following expansion in powers of m up to the fourth order

$$\begin{aligned} \Lambda^{-1}\mathcal{F}_z(m) &\approx \frac{Jz}{2}(c-2)m^2 - Bm + \frac{1}{2}\beta^{-1}(m^2 + \frac{1}{6}m^4) - \beta^{-1}\ln 2 \\ &= -k_B T \ln 2 + \frac{1}{2}k_B(T - T_c^L)m^2 + \frac{k_B T}{12}m^4 - Bm. \end{aligned} \tag{12.301}$$

Here $T_c^L = -k_B^{-1}(c-2)Jz = T_c$. For $m \to 0$ the only remaining term is $(-k_B T \ln 2)$, the entropy of the paramagnetic state [184].

Noting that $p_\pm = (1 \pm m)/2$ are the probabilities of finding a value of any s_i equal to ± 1, respectively, we recognize in $S(m)$ (12.299) the binary entropy. This may be expected for an independent-site model; in accordance with the general discussion in Section 4.8, the von Neumann (or Shannon) entropy as a function of order parameter (here m) is completely determined by the set of operators from which the Hamiltonian is constructed ($\hat{S}^z_{tot} = \sum_i \hat{\sigma}_i^z$ in the present case). On the other hand, the form of the internal energy function $U(m)$ in (12.300)

[66] On the other hand, Stirling formula may be used to justify the form of the binary Shannon entropy, [110], therefore it is implicitly present in our derivation.

is obtained simply by replacing operators in the MF Hamiltonians (either $\hat{H}_\lambda$, $\hat{H}_z$ or $\hat{H}$) by their expectation values. This is possible due to the specific form of $\hat{H}$ (12.283) (again, please consult Eq. (4.78) and a subsequent discussion of this point is Sec. (4.8)). Note, that the limit $m \to \pm 1$ of (12.300) is well-defined, and then $\mathcal{F}_z(m) \to U(m)$.Using $\hat{H}(m) = \hat{H}_\lambda(m, \lambda(m) = 0)$, one may define $\mathcal{F}_n(m)$ (4.46),

$$\begin{aligned} \mathcal{F}_n(m) &= \Lambda^{-1}\mathcal{F}(m, \lambda = 0) = -\beta^{-1} \ln \big(\mathrm{Tr}[\exp(-\beta \hat{H}_\lambda(m, \lambda(m) = 0))]\big) \\ &= \frac{Jz}{2} c m^2 - \beta^{-1} \ln[2 \cosh \big(\beta(Jzm + B)\big)]. \end{aligned} \tag{12.302}$$

Clearly, (12.300) and (12.302) are not identical. Although $\mathcal{F}(m, \lambda = 0)$ may be used to obtain the equilibrium value of magnetization (provided that we first put $c = 1$ in the Hamiltonian (12.296) to give it the HF form (4.109)), $\mathcal{F}_n(m)$ (12.302) has no reasonable interpretation for $m \neq m_0$. Consequently, $\mathcal{F}(m, \lambda = 0)$ cannot be consistently identified with the Landau potential, as discussed in Subsection 4.5.

12.3.4 Limit of zero temperature

In the $\beta \to \infty$ ($T \to 0$) limit, instead of looking for a minimum of $\mathcal{F}_z(m)$, one may directly minimize the ground-state energy, $U(m) = \Lambda u(m) \equiv E_G(m)$, i.e., the expectation value of the MF Hamiltonian

$$E_G(m) = \langle \hat{H}_\lambda \rangle = \langle \hat{H}_z \rangle_z = \langle \hat{H} \rangle = \Lambda \big(\frac{1}{2} Jz(c-2)m^2 - Bm\big). \tag{12.303}$$

The form of $E_G(m)$ may be inferred from Eq. (12.285), by assuming that the self-consistency conditions are fulfilled, and then when computing the average, the operator $\hat{S}^z_{tot}$ is simply replaced by its expectation value $\langle \hat{S}^z_{tot} \rangle = \Lambda m$. Also, because $[\hat{S}^z_{tot}, \hat{H}_\lambda] = 0$, Λm must be an eigenvalue of $\hat{S}^z_{tot}$. Consequently, for $B = 0$ and finite, even Λ, the permitted values of m are

$$m = \frac{\pm k}{\Lambda}, \qquad k = -\Lambda, -\Lambda + 2, \ldots, \Lambda \tag{12.304}$$

regardless the value of the molecular field λ. Note, that $E_G(m)$ (12.303) is well-defined for $m = \pm 1$. Alternatively, for $m \neq \pm 1$, (12.303) may be obtained as $\beta \to \infty$ limit of (12.300). The same limit applied to (12.295) yields the Hamiltonian, which in each subspace spanned by the eigenvectors of $\hat{S}^z_{tot}$ corresponding to the same eigenvalue Λm is proportional to the unit matrix.

For $B = 0$, $E_G(m)$ has two equivalent minima at $m = \pm 1$. Those minima lay at the boundary of the domain $\mathcal{D}_m = (-1, 1)$ and do not correspond to zeros of the derivative of $E_G(m)$, (in the case when that the discreteness of the allowed values of magnetization may be ignored and the calculus of variations can be applied in order to find extrema of $E_G(m)$). This is a simple example of a more general situation which may be encounter also in more complex MF models, as discussed in Supplement B (Subsection 12.2). Note, that within the finite-temperature treatment based on (12.300), for any nonzero T the analytical character of the minima at $m = m_0$, $|m_0| < 1$ is restored.

If we restrict ourselves to the pure $T = 0$ situation, we do not need to know the equilibrium value of $\lambda = \lambda_0$, as we are limited to the eigenstates of $\hat{H}$ (or equivalently, $\hat{H}_\lambda$). On the other hand, assuming continuous dependence of λ_0 on temperature, $\lambda_0(T)$, $\lambda_0(0)$ can be found by taking the limit of its finite-temperature value, given by (12.290)

$$\lambda_0(T = 0) = \lim_{T \to 0} \lambda_0(T) \equiv \lim_{m_0(T) \to 1} Jz(1 - c)m_0 = Jz(1 - c). \tag{12.305}$$

We explicitly assumed, that the symmetry is broken such that a state with $m_0 = 1$ is realized. Also, from (12.294) and (12.305) we infer that

$$\lim_{m_0(\beta)\to 1,\beta\to\infty} \beta^{-1}\text{arc tanh}(m_0) = Jz(2-c) + B. \tag{12.306}$$

12.3.5 Generalization to m-dependent exchange integral J

Within the present method we obtain self-consistent results for any form of the initial MF Hamiltonian $\hat{H}$, provided that the assumptions made in Subsection. 4.1 are fulfilled, i.e., $\hat{H}$ does not depend explicitly on time t. Consequently, within a given subclass (here the independent-site spin models), one may chose a particular MF Hamiltonian, which leads to results being in the best achievable agreement with either the experimental results or predictions of the exact models.

In the present case, let us consider the case of m - dependent exchange integral, $J = J(m)$ (for simplicity, full translational invariance has been assumed). We may also allow for the dependence of J on parameters $\vec{b}$, not being the expectation values of any operators, thus for which self-consistency conditions are no demanded. With such a modification, the form of the Hamiltonian $\hat{H}(m)$ (12.285), $\lambda(m)$ (12.294), $\hat{H}_z(m)$ (12.295) as well as that of free energy (12.287) remain the same (obviously, with $J \to J(m) \equiv J$). However, the equations (12.289) and (12.292) are modified according to

$$(Jzcm + \frac{1}{2}J'zcm^2 + \lambda) - z(J + J'm)\tanh\big(\beta(Jzm + B + \lambda)\big) = 0, \tag{12.307}$$

$$\lambda_0 = -\Big\langle \frac{\partial \hat{H}}{\partial m} \Big\rangle_{m=m_0} = Jz(1-c)m_0 + \frac{1}{2}J'z(2-c)m_0^2, \tag{12.308}$$

with $J' = dJ/dm$. The equilibrium value of magnetization is now given by

$$m_0 = \tanh\Big(\beta\big(z(2-c)(Jm_0 + \frac{1}{2}J'm_0^2) + B\big)\Big). \tag{12.309}$$

Different forms of $J(m)$ yield in general different values of transition temperature T_c as well as other characteristics of the system. We will not discuss the choice of $J(m)$ that leads to various desired properties of the model, nor analyze the general requirements that $J(m)$ should satisfy as a function of m. Instead, below we quote an example. In Ref. [189], devoted to the problem of magnetic ordering in rare-earth metals, the Bragg-Williams approximation is used, with

$$J(m) = m^2\Big(B_0 + 2B_1\cos\alpha + 2(B_2 - Cm^2\cos(2\alpha))\Big). \tag{12.310}$$

In above, B_0, B_1, B_2 are m-independent constants, and $b_1 = \alpha$ is an additional variational parameter. Note, that when solving such MF model, in contrast to Ref. [189] we are not forced to drop out the quartic term in (12.310). In fact, within our approach, arbitrary form of $J(m)$ can be consistently analyzed.

12.3.6 General solution with non-uniform magnetization

We now relax the simplifying assumption of the full translational symmetry, which has been used to derive the Hamiltonian (12.285). We search for the solution characterized by the site-dependent magnitude of magnetization, but with the same quantization axis (z axis) for each site. Instead of a single scalar order parameter m we have now Λ such variables, $m_1, m_2, \ldots m_\Lambda$, the same number of the Lagrange multipliers, $\lambda_1, \lambda_2, \ldots \lambda_\Lambda$, and values of the external magnetic

field, $B_1, B_2, \ldots B_\Lambda$. Most steps proceed analogously as in the homogeneous case. Instead of (12.285), the MF Hamiltonian augmented with the constraints reads now

$$\begin{aligned}\hat{H}_\lambda &= \sum_{\langle ij\rangle} J_{ij}\big(\hat{\sigma}_i^z m_j + \hat{\sigma}_j^z m_i - c m_i m_j\big) - \sum_i B_i \hat{\sigma}_i^z - \sum_i \lambda_i\big(\hat{\sigma}_i^z - m_i\big) \\ &= \sum_i \big(\sum_j J_{ij} m_j - B_i - \lambda_i\big)\hat{\sigma}_i^z + \sum_i \big(\lambda_i m_i - \frac{c}{2}\sum_j m_i m_j\big). \end{aligned} \tag{12.311}$$

Having (12.311) at our disposal, we can easily find the explicit form of the generalized Landau potential, $\mathcal{F} = \mathcal{F}(m_1, \ldots, m_\Lambda, \lambda_1, \ldots, \lambda_\Lambda; \beta, B_1, \ldots, B_\Lambda)$,

$$\mathcal{F} = -\frac{1}{\beta}\sum_i \ln\big[2\cosh\Big(\beta\big(\sum_j J_{ij} m_j - B_i - \lambda_i\big)\Big)\big] + \sum_i \lambda_i m_i - \frac{c}{2}\sum_{ij} m_i m_j \tag{12.312}$$

Equating the derivative with respect to λ_k to zero, we obtain

$$m_k + \tanh\Big(\beta\big(\sum_j J_{kj} m_j - B_k - \lambda_k\big)\Big) = 0, \tag{12.313}$$

which may be easily solved for λ_k. Inserting the result back into (12.312) and making again use of identities (12.297) we obtain Landau free energy potential

$$\mathcal{F}_z = \Big(1 - \frac{c}{2}\Big)\sum_{ij} m_i J_{ij} m_j - \sum_i B_i m_i - \sum_i \beta^{-1}\mathrm{S}(m_i). \tag{12.314}$$

Similarly to the $m_i = m$ case, $\mathrm{S}(m_i)$ appearing in (12.314) has the form of the Shannon entropy,

$$\mathrm{S}(m_i) = \frac{1 - m_i}{2}\ln\Big(\frac{1 - m_i}{2}\Big) + \frac{1 + m_i}{2}\ln\Big(\frac{1 + m_i}{2}\Big). \tag{12.315}$$

Again, this particular form of $\mathrm{S}(m_i)$ results from the single-site character of operators $\hat{S}_i^z$ appearing in $\hat{H}_\lambda$ (12.311).

By equating the derivatives of (12.314) with respect to m_k to zero, we obtain the set of coupled, nonlinear equations, which are intractable in general. Putting $J_{ij} = -J$, for nearest neighbors and $J_{ij} = 0$ otherwise, we can rewrite the first term of (12.314) as

$$\sum_{ij} m_i J_{ij} m_j = -J\sum_{ij} m_i m_j = -J\big(\sum_i m_i(\Delta m)_i + z m_i^2\big). \tag{12.316}$$

Δ is the discrete version of $z/2 = d$-dimensional Laplace operator,

$$(\Delta m)_i = \sum_{\nu=1}^{d} \frac{m(i + a\hat{x}_\nu) - 2m(i) + m(i - a\hat{x}_\nu)}{a^2}, \tag{12.317}$$

where $m(i) \equiv m_i$, and a denotes the lattice constant. Using the above definitions, and expanding $\mathrm{S}(m_i)$ up to fourth power in m_i, we obtain

$$\mathcal{F}_{GL} = \frac{J}{2}(c - 2)\sum_i \big(m_i(\Delta m)_i + z m_i^2\big) + k_B T\sum_i \big(\frac{1}{2}m_i^2 + \frac{1}{12}m_i^4\big) - \sum_i B_i m_i. \tag{12.318}$$

It the limit $\Lambda \to \infty$, $a \to 0$, and $a\Lambda = \text{const}$, magnetization becomes a continuous field, $m_i \to \phi(x), x \in \mathbb{R}^d$. Then (12.314) or (12.318) (after integrating by parts and dropping the surface terms) becomes the Ginzburg-Landau (G-L) functional for the case of one real, scalar order parameter field $\phi(x)$. Strictly speaking, this limit lies beyond the scope of our method. This is due to the assumption we have made previously, that a dimension D_H of the Hilbert space is finite (here $D_H = 2^\Lambda$). Nonetheless, for the MF model defined by the Hamiltonian (12.311), formula (12.318) gives the correct discrete form of the G-L functional.

12.3.7 Quantum fluctuations

As an illustration to the general considerations of Supplement A (Subsection 12.1), let us estimate the magnitude of quantum fluctuations of an order parameter (magnetization). This may be done without solving the model in the most general case, i.e., without a detailed knowledge of the equilibrium configuration $(m_{10}, m_{20}, \ldots m_{\Lambda 0})$. Clearly, at low temperatures, and for $J_{ij} = -J < 0$ we expect uniform magnetic state with $m_{i0} = m_0 \approx 1$ to minimize (12.314). For $\hat{A}_s = \hat{\sigma}_i^z$, the r.h.s. of Eq. (12.268) reads

$$\sigma_r^2(\hat{\sigma}_i^z) = \frac{\mathrm{var}(\hat{\sigma}_i^z)}{(m_i)^2} = \frac{1 - m_i^2}{m_i^2}. \tag{12.319}$$

For $m_i \approx 1$, $\sigma_r^2(\hat{\sigma}_i^z) \approx 0$, whereas for $m_i \to 0$, $\sigma_r^2(\hat{\sigma}_i^z) \to \infty$. Also, for $m_i = 0.99, 0.95, 0.9, 0.5$ we obtain, respectively $\sigma_r^2(\hat{\sigma}_i^z) = 0.02, 0.11, 0.23, 3$. Even for relatively small deviations of each m_i from $m_{i0} \approx 1$, quantum fluctuations are rather pronounced. Also, the inequality $1 - m_i \leq \sqrt{1 - m_i^2}$ holds for $0 \leq m_i \leq 1$, therefore the quantum dispersion exceeds the difference between a given m and the maximal possible value $m_{i0} = 1$. Consequently, uncertainty of m_i due to 'quantum' fluctuations ($\sim \sqrt{1 - m_i^2}$) is always greater then its departure from the equilibrium value ($\sim (1 - m_i)$).

The above analysis indicates, that MF approach for the Ising model in the form defined by $\hat{H}_\lambda$ (12.311), is internally consistent at temperatures, for which the equilibrium mean-field configuration remains close to the saturated configuration with $m_{i0} = m_0 \approx 1$. Only in such a situation, both the quantum and classical fluctuations could be neglected. If we move towards the transition (Curie) temperature T_c, fluctuations become so pronounced, that the mean-field picture loses its validity.

References

[1] *Handbook of High-Temperature Superconductivity*, edited by J. R. Schrieffer and J. S. Brooks, Springer Science, New York (2007).

[2] E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994).

[3] N. P. Armitage, P. Fournier, and R. L. Greene, Rev. Mod. Phys. **82**, 2010.

[4] L. Taillefer, *Scattering and Pairing in Cuprate Superconductors*, Annu. Rev. Condens. Matter Phys. 2010.1:51-70.

[5] M. R. Norman and C. Pépin, *The electronic nature of high temperature cuprate superconductors*, Rep. Prog. Phys. **66**, 1547–1610 (2003).

[6] M. R. Norman, D. Pines, and C. Kallin, *The pseudogap: friend or foe of high T_c?*, Advances in Physics, Vol. 54, No. 8, 715–733 , December 2005.

[7] A. J. Leggett, Nature Physics, vol. 2, **134** (2006).

[8] M. Ogata and H. Fukuyama, Rep. Prog. Phys. **71**, 036501 (2008) and References therein.

[9] J. G. Bednorz and K. A. Müller, Z. Phys. B **64**, 189; J. G. Bednorz and K. A. Müller, Rev. Mod. Phys. **60**, 585.

[10] D. Vaknin, S. K. Sinha, D. E. Moncton, D. C. Johnston, J. M. Newsam, C. R. Safinya, and H. E. King, Jr., Phys. Rev. Lett. **58**, 2802 (1987).

[11] M. A. Kastner and R. J. Birgenau, Rev. Mod. Phys. **70**, 897 (1998).

[12] C. C. Tsuei and J. R. Kirtley, Rev. Mod. Phys. **72**, 969.

[13] B. Keimer, N. Belk, R. J. Birgeneau, A. Cassanho, C. Y. Chen, M. Greven, M. A. Kastner, A. Aharony, Y. Endoh, R. W. Erwin, and G. Shirane, Phys Rev. B **46** 14034 (1992).

[14] T. Nakano, N. Momono, and M. Ido, J. Phys. Soc. Jpn. **67**, 2622, (1998).

[15] A. Fujimori et al., in *Physics and Chemistry of Transition Metal Oxides*, edited by H. Fukuyama and M. Nagaosa (Springer Verlag, Berlin, 1999) pp. 111ff.

[16] G. A. Thomas, J. Orenstein, D. H. Rapkine, M. Capizzi, A. J. Millis, R. N. Bhatt, L. F. Schneemeyer, and J. V. Waszczak, Phys. Rev. Lett. **61**, 1313 (1988).

[17] H. Alloul, T. Ohno, and P. Mendels, Phys. Rev. Lett. **63**, 1700 (1989).

[18] Ch. Renner, B. Revaz, J.-Y. Genoud, K. Kadowaki, and O. Fischer, Phys. Rev. Lett. **80**, 149 (1998).

[19] J. Hubbard, Proc. R. Soc. London, Ser. A **276**, 238 (1963); ibid. **281**, 401 (1964).

[20] M. C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963); Phys. Rev. **134**, A923 (1964); Phys. Rev. **137**, A1726 (1965).

[21] J. Kanamori, Prog. Theor. Phys. **30**, 275 (1963).

[22] J. Spałek, A. M. Oleś, Jagiellonian University preprint SSPJU-6/76, (1976).

[23] K. A. Chao, J. Spałek, A. M. Oleś, J. Phys. C **10**, L271 (1977).

[24] K. A. Chao, J. Spałek, A. M. Oleś, Phys. Rev. B **18**, 3453 (1978).

[25] A. H. MacDonald, S. M. Girvin, and D. Yoshioka, Phys. Rev. B **37**, 9753 (1988).

[26] P. W. Anderson, in: *Frontiers and Borderlines in Many-Particle Physics*, edited by R.A. Broglia and J. R. Schrieffer (North-Holland, Amsterdam, 1988) p.1 ff.

[27] J. Spałek, Phys. Rev. B **37**, 533 (1988).

[28] For historical overview see: J. Spałek, Acta Phys. Polon. A **111**, 409 (2007) and References therein.

[29] F. C. Zhang and T. M. Rice, Phys. Rev. B **37**, 3759 (1988).

[30] P. A. Lee, N. Nagaosa and X-G Wen, Rev. Mod. Phys **78**, 17 (2006) and References therein.

[31] P. A. Lee, *From high temperature supercondutivity to quantum spin liquid: progress in strong correlation physics*, arXiv:0708.2115v2 [cond-mat.str-el]

[32] Y. Hasegawa and D. Poilblanc, Phys. Rev. B **40**, 9035 (1989); cf. also: D. Poilblanc, arXiv:cond-mat/9505004v1.

[33] B. Edegger, V. N. Muthukumar, and C. Gros, Adv. Phys. **56**, 927 (2007) and references therein.

[34] L. Pauling, Phys. Rev. **54**, 899 (1938).

[35] L. Pauling, Nature **161**, 1019 (1948).

[36] P. W. Anderson, Mat. Res. Bull **8**, 153 (1973).

[37] P. Fazekas and P. W. Anderson, Philos. Mag. **30**, 432 (1973).

[38] P. W. Anderson, Science **235**, 1196 (1987).

[39] P. W. Anderson, P. A. Lee, M. Randeria, T. M. Rice, N. Trivedi and F. C. Zhang, J. Phys.: Condens. Matter **16**, (2004) R755-R769.

[40] P. W. Anderson, Phys. Rev. Lett. **96**, 017001 (2006).

[41] J. Bardeen, L. N. Cooper, J. R. Schrieffer, Phys. Rev. **106**, 162, (1957); ibid., **108**, 1175, (1957).

[42] M. Tinkham, *Introduction to Superconductivity*, 2nd edition, Dover (2004).

[43] G. Rickayzen, *Theory of superconductivity* (J. Wiley and Sons, New York, 1965).

[44] J. R. Schrieffer, *Theory of Superconductivity*, Advanced Books Classics.

[45] C. Gros, Phys. Rev. B **38**, 931 (1988); Ann. Phys. **189**, 53 (1989).

[46] H. Yokoyama and H. Shiba, J. Phys. Soc. Jpn. **57**, 2482 (1988).

[47] A. Himeda and M. Ogata, Phys. Rev. Lett. **85**, 4345 (2000).

[48] A. Paramekanti, M. Randeria, and N. Trivedi, Phys. Rev. Lett. **87**, 217002 (2001); Phys. Rev. B **70**, 054504 (2004).

[49] S. Sorella, G. B. Martins, F. Becca, C. Gazza, L. Capriotti, A. Parola, and E. Dagotto, Phys. Rev. Lett. **88**, 117002 (2002).

[50] C. T. Shih, T.K. Lee, R. Eder, C.-Y. Mou, and Y. C. Chen, Phys. Rev. Lett. **92**, 227002 (2004).

[51] D. A. Ivanov, Phys. Rev. B **70**, 104503 (2004).

[52] M. Lugas, L. Spanu, F. Becca, and S. Sorella Phys. Rev. B **74**, 165122 (2006).

[53] L. Spanu, M. Lugas, F. Becca, and S. Sorella, Phys. Rev. B **77**, 024510 (2008).

[54] Manuela Capello, Marcin Raczkowski, and Didier Poilblanc, Phys. Rev. B **77**, 224502 (2008).

[55] Chung-Pin Chou, N. Fukushima, and Ting Kuo Lee, Phys. Rev. B **78**, 134530 (2008).

[56] S. Pathak, V. B. Shenoy, M. Randeria, and N. Trivedi, Phys. Rev. Lett. **102**, 027002 (2009).

[57] L. Goren and E. Altman, Phys. Rev. Lett. **104**, 257002 (2010).

[58] F. C. Zhang, C. Gros, T. M. Rice, and H. Shiba, Supercond. Sci. Technol. **1**, 36 (1988) [preprint at: arXiv:cond-mat/0311604v1].

[59] D. Vollhardt, Rev. Mod. Phys. **56**, 99, (1984).

[60] W.-H. Ko, C. P. Nave, and P. A. Lee, , Phys. Rev. B **76**, 245113 (2007).

[61] B. Edegger, V. N. Muthukumar, C. Gros, and P. W. Anderson, Phys. Rev. Lett. **96**, 207002 (2006).

[62] F. C. Zhang, Phys. Rev. Lett. **64**, 974, (1990).

[63] M. Sigrist, T. M. Rice and F.C. Zhang, Phys. Rev. B **49**, 12 058 (1994).

[64] H. Tsuchiura, Y. Tanaka, M. Ogata, and S. Kashiwaya, J. Phys. Soc. Jpn. **68**, 2510 (1999).

[65] H. Tsuchiura, Y. Tanaka, M. Ogata, and S. Kashiwaya, Phys. Rev. Lett. **84**, 3165 (2000).

[66] M. Ogata and A. Himeda, J. Phys. Soc. Jpn. **72**, 374, (2003).

[67] M. Ogata, J. Phys. Soc. Jpn. **72**, 1839 (2003).

[68] D. Poilblanc, Phys. Rev. B **72**, 060508(R), (2005).

[69] J. Bünemann, F. Gebhard, K. Radnóczi, and P. Fazekas, J. Phys.: Condens. Matter **17**, 3807 (2005).

[70] C. Li, S. Zhou and Z. Wang, Phys. Rev. B **73**, 060501(R) (2006).

[71] Qiang-Hua Wang, Z. D. Wang, Yan Chen, and F. C. Zhang, Rev. B **73** 092507 (2006).

[72] C. Weber, D. Poilblanc, S. Capponi, F. Mila, and C. Jaudet, Phys. Rev. B **74**, 104506 (2006).

[73] M. Raczkowski, D. Poilblanc, R. Frésard, and A. M. Oleś, Phys. Rev. B **75**, 094505 (2007).

[74] M. Raczkowski, M. Capello, D. Poilblanc, R. Frésard, and A. M. Oleś, Phys. Rev. B **76**, 140505(R) (2007).

[75] M. Raczkowski, Acta Phys. Polon. A **114**, 107 (2008).

[76] N. Fukushima, Phys. Rev. B **78**, 115105 (2008).

[77] L. R. Niestemski and Ziqiang Wang, Phys. Rev. Lett. **102**, 107001 (2009).

[78] M. Raczkowski, D. Poilblanc, Phys. Rev. Lett. **103**, 027001 (2009).

[79] K.-Y. Yang, Wei Qiang Chen, T. M. Rice, M. Sigrist and F.-C. Zhang, New Journal of Physics **11**, 055053 (2009).

[80] K.-Y. Yang, C. T. Shih, C. P. Chou, Huang S. M., T. K. Lee, T Xiang and F-C. Zhang, Phys. Rev. B **73** 224513, (2009).

[81] N. Fukushima, Chung-Pin Chou, and Ting Kuo Lee, Phys. Rev. B **79**, 184510 (2009).

[82] J. Jędrak and J. Spałek, Phys. Rev. B **81**, 073108 (2010), also as arXiv:0908.4411.

[83] J. Jędrak and J. Spałek, Physical Review B **83**, 104512 (2011); also as arXiv:1004.4282.

[84] J. Jędrak and J. Spałek, arXiv:1002.1607 (unpublished).

[85] J. Bünemann, F. Gebhard, and R. Thul, Phys. Rev. B **67**, 075103 (2003).

[86] G. Baskaran, arXiv:0611548.

[87] C. Gros, B. Edegger, V. N. Muthukumar, and P. W. Anderson, PNAS 103, 14298-14301 (2006); also as arXiv:0606.750.

[88] J.C. Campuzanno, H. Ding, M. R. Norman, H. M. Fretwell, M. Randeria, A. Kaminski, J. Mesot, T. Takeuchi, T. Sato, T. Yokoya, T. Takahashi, T. Mochiku, K. Kadowaki, P. Guptasarma, D. G. Hinks, Z. Konstantinovic, Z. Z. Li, and H. Raffy, Phys. Rev. Lett. **83**, 3709 (1999).

[89] A. Kaminski, M. Randeria, J. C. Campuzano, M. R. Norman, H. Fretwell, J. Mesot, T. Sato, T. Takahashi, and K. Kadowaki, Phys. Rev. Lett. **86**, 1070 (2001).

[90] A. A. Kordyuk, S. V. Borisenko, A. Koitzsch, J. Fink, M. Knupfer, and H. Berger, Phys. Rev. B **71**, 214513 (2005).

[91] A. Kaminski, S. Rosenkranz, H. M. Fretwell, M. R. Norman, M. Randeria, J. C. Campuzano, J-M. Park, Z. Z. Li, and H. Raffy, Phys. Rev. B **73**, 174511 (2006).

[92] T. Yoshida, X. J. Zhou, K. Tanaka, W. L. Yang, Z. Hussain, Z.-X. Shen, A. Fujimori, S. Sahrakorpi, M. Lindroos, R. S. Markiewicz, A. Bansil, Seiki Komiya, Yoichi Ando, H. Eisaki, T. Kakeshita, and S. Uchida, Phys. Rev. B **74**, 224510 (2006).

[93] S. V. Borisenko, A. A. Kordyuk, V. Zabolotnyy, J. Geck, D. Inosov, A. Koitzsch, J. Fink, M. Knupfer, B. Büchner, V. Hinkov, C. T. Lin, B. Keimer, T. Wolf, S. G. Chiuzbaian, L. Patthey, and R. Follath, Phys. Rev. Lett. **96**, 117004 (2006).

[94] G. Baskaran, Z. Zou, and P. W. Anderson, Solid State Commun. B **63**, 973 (1987).

[95] P. W. Anderson, G. Baskaran, T. Hsu and Z. Zou, Phys. Rev. Lett. **58**, 2790 (1987).

[96] G. Baskaran and P. W. Anderson, Phys. Rev. B **37**, 580 (1988).

[97] G. Kotliar and J. Liu, Phys. Rev. B **38**, 5142 (1988).

[98] Y. Suzumura, Y. Hasegawa, and H. Fukuyama, J. Phys. Soc. Jpn. **57**, 2768, (1988).

[99] F. Gebhard, Phys. Rev. B **41**, 9452 (1990).

[100] F. Gebhard, Phys. Rev. B **44**, 992 (1991).

[101] J. Bünemann and F. Gebhard, Phys. Rev. B **76**, 193104 (2007).

[102] J. Jędrak, J Kaczmarczyk and J. Spałek, arXiv:1008.0021 (unpublished).

[103] G. Kotliar and A. N. Ruckenstein, Phys. Rev. Lett **57**, 1362 (1986).

[104] X.-G. Wen and P. A. Lee, Phys. Rev. Lett. **76**, 503 (1996).

[105] P. A. Lee, N. Nagaosa, T.-K. Ng, and X.-G. Wen, Phys. Rev. B. **57**, 6003 (1998).

[106] P. A. Lee, N. Nagaosa, and X-G. Wen, Rev. Mod. Phys. **78**, 17 (2006).

[107] R. P. Feynman *Statistical Mechanics. A Set of Lectures*, W. A. Benjamin, (1972).

[108] A. Decoster, J. Phys. A: Math. Gen. **37** 9051-9070 (2004).

[109] E. T. Jaynes, Phys. Rev. **106**, 620-630 (1957); **108**, 171-190 (1957).

[110] E. T. Jaynes, IEEE Trans. on Systems Science and Cybernetics SSC-4, 227, (1968).

[111] E. T. Jaynes, *Probability Theory: The Logic of Science*, Cambridge University Press, (2003).

[112] L. E. Reichl, *A Modern Course in Statistical Physics*, 2nd Edition, John Wiley and Sons Inc., New York (1998).

[113] J. Jędrak and J. Spałek, arXiv:0804.1376 (unpublished).

[114] I. Affleck and J. B. Marston, Phys. Rev. B **37**, 3774, (1988); Phys. Rev. B **39**, 11538, (1989).

[115] P. Lederer, D. Poilblanc, and T. M. Rice, Phys. Rev. Lett. **63**, 1519 (1989).

[116] T. C. Hsu, Phys. Rev. Lett. **41**, 11379 (1990).

[117] H. Shiba and M. Ogata, J. Phys. Soc. Jpn. **59**, 2971, (1990).

[118] T. C. Hsu, I. Affleck and J. B. Marston, Phys. Rev. B **43**, 2866 (1991).

[119] M. U. Ubbens and P. A. Lee, Phys. Rev. B **46**, 8434 (1992).

[120] T.-H. Gimm and S.-H. Suck Salk, Phys. Rev. B **62**, 13930 (2000).

[121] M. Andersson and S. Östlund, Phys. Rev. B **67**, 014403 (2003).

[122] K. Hamada and D. Yoshioka, Phys. Rev. B **67**, 184503 (2003).

[123] H. Zhao and J. Engelbrecht, Phys. Rev. B **71**, 054508 (2005).

[124] D. Bensimon and N. Nagaosa, Phys. Rev. B **73**, 184517 (2006).

[125] H. Yamase and H. Kohno, J. Phys. Soc. Jpn. **69**, 332 (2000); *ibid.* **69**, 2151, (2000).

[126] L. Dell'Anna and W. Metzner, Phys. Rev. B **73**, 045127 (2006).

[127] A. Miyanaga and H. Yamase , Phys. Rev. B **73**, 174513 (2006).

[128] B. Edegger, V. N. Muthukumar, and C. Gros, Phys. Rev. B **74**, 165109 (2006).

[129] H. Yamase and W. Metzner, Phys. Rev. B **75**, 155117 (2007).

[130] H. Yamase, arXiv:0907.4253 (2009).

[131] J. D. Van der Waals, *Over de Continuiteit van den Gas en Vloeistoftoestand*, thesis, Leiden (1873).

[132] Leo P. Kadanoff, *Theories of Matter: Infinities and Renormalization*, arXiv:1002.2985.

[133] J. J. Binney, N. J. Dowrick, A. J. Fisher and M. E. J. Newman, *The theory of Critical Phenomena. An Introduction to the Renormalization Group*. Clarendon Press, Oxford, 1992. Polish edition: *Zjawiska krytyczne. Wstęp do teorii grupy renormalizacji*, Państwowe Wydawnictwo Naukowe PWN (1998).

[134] P. Weiss, J Phys. **6**, 661, (1907).

[135] C. Kittel, *Introduction to Solid State Physics*, fourth edition, John Wiley and Sons (1971).

[136] H. Bruus, K. Flensberg, *Many-Body Quantum Theory in Condensed Matter Physics. An Introduction*, Oxford University Press, (2004).

[137] P. Ring and P. Schuck, *The Nuclear Many-body Problem* (Springer-Verlag, New York, 1980).

[138] J. Ginter, *Wstęp do fizyki atomu, cząsteczki i ciała stałego*, Państwowe Wydawnictwo Naukowe PWN (1979) (in polish).

[139] Y. Nambu and G. Jona-Lasino, Phys. Rev. **122**, 345, (1961).

[140] R. Balian, J. M. Drouffe, and C. Itzykson, Phys. Rev. D **10**, 3376 (1974).

[141] S. Elitzur, Phys. Rev. D **12**, 3978 (1975).

[142] A. Bohr, B. R. Mottelson and D. Pines, Phys. Rev. **110**, 936 (1958).

[143] B. Carter, N. Chamel, and P. Haensel, *Effect of BCS pairing on entrainment in neutron superfluid current in neutron star crust*, Nuclear Physics A **759**P (2005) 441-464.

[144] I. N. Levine, *Quantum Chemistry*, fifth edition, Prenitce Hall, (1991).

[145] P. W. Anderson, *More is Different*, Science 177 (4047): 393-396.

[146] Leo P. Kadanoff, *More is the Same; Phase Transitions and Mean Field Theories*, arXiv:0906.0653.

[147] T. W. B. Kibble, *Symmetry Breaking and Defects*, in: *Patterns of Symmetry Breaking*, NATO Science Series, Series II, Vol. 127, p. 3, edited by H. Arodź, J. Dziarmaga and W. H. Zurek.

[148] L. D. Landau, I. M. Lifshitz, Pitaevskii, *Statistical Physics*, Dover. Polish edition: *Termodynamika Statystyczna*, Państwowe Wydawnictwo Naukowe PWN (1974).

[149] L. D. Landau, Phys. Z. Sow. **11** 26 545 (1937). English Translation: *Collected Papers of L. D. Landau* edited by D. ter Haar (Oxford, Pergamon Press, 1965) pp 193-215.

[150] L. D. Landau, Sov. Phys. JETP **3**, 920 (1957); *ibidem*, **5**, 101 (1957).

[151] V. L. Ginzburg, L. D. Landau, JETP **20**, 1064 (1950).

[152] W. L. Bragg and E. J. Williams, Proc. Roy. Soc. London 145 **A**, 699 (1934).

[153] M. Schlosshauer, *Decoherence, the measurement problem, and interpretations of quantum mechanics*, arXiv:quant-ph/0312059.

[154] L. E. Ballentine, Rev. Mod. Phys., **42**, 358 (1970).

[155] N. Wiebe, L. E. Ballentine, Phys. Rev. A **72**, 022109 (2005).

[156] L. E. Ballentine, Found. Phys. **38**, 916 (2008).

[157] J. Zaanen, *The Classical Condensates: From Crystals to Fermi-liquids*, Leiden University, Leiden (1996).

[158] J. Zaanen, A. M. Oles, Ann. Physik 5, p. 224 (1996).

[159] R. P. Feynman, A. R. Hibbs *Quantum Mechanics and Path Integrals*, McGraw-Hill, (1965).

[160] J. W. Negele, H. Orland, *Quantum Many-particle Systems*.

[161] E. Fradkin, *Field Theories of Condensed Matter Systems*, Addison-Wesley (1991).

[162] R. Rossignoli and A. Plastino, Phys. Rev. C **40**, 1798 (1989).

[163] R. Rossignoli and A. Plastino, Phys. Rev. A **42**, 2065 (1990).

[164] R. Rossignoli, A. Plastino, H.G. Miller, Phys. Rev. C **43**, 1599 (1991).

[165] M. Gaudin, Nucl. Phys. **15**, 89 (1960).

[166] S. Sachdev, *Quantum Phase Transitions*, Cambridge Univeristy Press (1999).

[167] E. Ising, Z. Phys. **31** 253 (1925).

[168] N. N. Bogoliubov, *Quasiaverages in Problems of Statistical Mechanics*. Communication JINR, D-781, JINR, Dubna (1961); *On the Principle of the Weakening of Correlations in the Method of Quasiaverages*, *ibidem*, P-549, JINR, Dubna (1961) (in Russian).

[169] A. L. Kuzemsky, arXiv:1003.1363.

[170] Herbert Wagner, Zeitschrift für Physik 195, 273-299 (1966).

[171] P. C. Hohenberg, Phys. Rev. **158**, 383 (1967).

[172] N. D. Mermin and H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966).

[173] Gang Su, A. Schadschneider, J. Zittartz, Physics Letters A 230, (1997), 99-104.

[174] Gang Su and Masao Suzuki, Phys. Rev. B **58**, 117 (1998).

[175] Gang Su, Phys. Rev. Lett. **86**, 3690 (2001).

[176] Gang Su, Phys. Rev. B **72**, 092510 (2005).

[177] R. Balian and M. Veneroni, Annals of Physics **187**, 29-78 (1988).

[178] D. N. Zubarev, *Statistical Mechanics of Non-Equilibrium Processes*, (In Russian, Izd. Nauka, Moskva 1964); Polish edition: *Termodynamika Statystyczna*, Państwowe Wydawnictwo Naukowe PWN (1974).

[179] R. K. Niven, World Scientific, Vol. 9, 2009, pp 283-318, arXiv:0907.2732.

[180] R. P. Feynman, Phys. Rev. **56**, 340 (1939).

[181] R. Jastrow, Phys. Rev. **98**, 1479 (1955).

[182] J. Kaczmarczyk and J. Spałek, Phys. Rev. B **79**, 214519 (2009).

[183] M. M. Maśka, M. Mierzejewski, J. Kaczmarczyk, and J. Spałek, Phys. Rev. B **82**, 054509 (2010).

[184] K. Huang, *Statistical Mechanics*, Polish edition, *Fizyka Statystyczna*, Państwowe Wydawnictwo Naukowe PWN, Warszawa (1973).

[185] P. N. Argyres, T. A. Kaplan, N. P. Silva, Phys. Rev. A **9**, 1716 (1974).

[186] M. A. Nielsen, I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, (2000).

[187] C. Beck, arXiv:0902.1235.

[188] C. Tsallis, J. Stat. Phys. **52**, 479 (1988); C. Tsallis, U. Tirnakli, arXiv:0911.1263.

[189] R. J. Elliott, Phys. Rev. **124**, 346 (1961).

[190] P. Fazekas, *Lecture Notes on Electron Correlations and Magnetism*, World Scientific, Singapore (1999).

[191] A. L. Chernyshev, D. Galanakis, P. Philips, A. V. Rozhkov, and A.-M. S. Tremblay, Phys. Rev. B **70**, 235111 (2004).

[192] J.-Y. Delannoy, M. J. P. Gingras, P. C. W. Holdsworth, and A.-M. S. Tremblay, Phys. Rev. B **72**, 115114 (2005).

[193] R. Coldea, S. M. Hayden, G. Aeppli, T. G. Perring, C. D. Frost, T. E. Mason, S.-W. Cheong, and Z. Fisk, Phys. Rev. Lett. **86**, 5377 (2001).

[194] G. J. Chen, R. Joynt, and F. C. Zhang, J. Phys.: Condens. Matter **3**, 5213 (1991).

[195] M.M. Korshunov and S.G. Ovchinnikov, Eur. Phys. J. B **57**, 271 (2007).

[196] L. L. Foldy and S. A. Wouthuysen, Phys. Rev. **78**, 29 (1950).

[197] P. W. Anderson, Phys. Rev. **115**, 2 (1959).

[198] P. W. Anderson, Solid State Physics, vol. 14, pp. 99-213 (1963).

[199] M. Raczkowski, Ph.D. thesis, Jagiellonian University, Kraków, 2005.

[200] F. C. Zhang, Phys. Rev. Lett. **90**, 207002 (2003).

[201] S. Guertler, Q.-H. Wang, and F. C. Zhang, Phys. Rev. B **79**, 144526 (2009).

[202] Wan-Shen Wang, Xiang-Mei He, Da Wang, and Qiang-Hua Wang, Z. D. Wang and F. C. Zhang, Phys. Rev. B **82**, 125105 (2010).

[203] J. Spałek, A. M. Oleś, and K. A. Chao, phys. stat. sol. (b) **108**, 329 (1981).

[204] J. Spałek, Condensed Matter Physics, Vol. 11, No 3(55), 455 (2008).

[205] W. Metzner and D. Vollhardt, Phys Rev Lett. **59**, 121 (1987); Phys Rev. B **37**, 7382 (1988).

[206] W. Metzner and D. Vollhardt, Phys Rev Lett. **62**, 324 (1989).

[207] J. Kurzyk, Ph.D. thesis, Jagiellonian University, Kraków, 2007.

[208] J. Bünemann, T. Schickling, and F. Gebhard, unpublished; J. Bünemann, private communication.

[209] J. Spałek, A. M. Oleś and J. M. Honig, Phys. Rev. B **28**, 6802 (1983); J. Spałek, A. Datta and J. M. Honig, Phys. Rev. B **33**, 4891 (1986).

[210] K. Seiler et al., J. Low. Temp. Phys. **64**, 195 (1986).

[211] J. Spałek, A. Datta and J. M. Honig, Phys. Rev. Lett. **59**, 728 (1987).

[212] J. Jędrak and J. Spałek, unpublished.

[213] M.S. Hybertsen, E. B. Stechel, M. Schluter, D. R. Jennison, Phys. Rev. B **41**, 11068 (1990).

[214] E. Pavarini, I. Dasgupta, T. Saha-Dasgupta, O. Jepsen, and O. K. Andersen, Phys. Rev. Lett. **87**, 047003 (2001).

[215] H. Tsuchiura, M. Ogata, Y. Tanaka, and S. Kashiwaya, Phys. Rev. B **68**, 012509 (2003).

[216] J. Kaczmarczyk and J. Spałek, unpublished.

[217] Author is indebted to Prof. A. M. Oleś and Prof. K. Rościszewski for pointig out this fact.

[218] T. Eckl, D. J. Scalapino, E. Arrigoni, and W. Hanke, Phys. Rev. B **66**, 140510(R) (2002).

[219] S. Banerjee, T. V. Ramakrishnan, and Ch. Dasgupta, Phys. Rev. **B** 83, 024510 (2011).

[220] B. Edegger, N. Fukushima, C. Gros, and V. N. Muthukumar, Phys. Rev. B **72**, 134504 (2005).

[221] N. Fukushima, B. Edegger, V. N. Muthukumar, and C. Gros, Phys. Rev. B **72**, 144505 (2005).

[222] M. A. Leontovich, *Statistical Physics*. Polish edition: *Fizyka Statystyczna*, Państwowe Wydawnictwo Naukowe PWN, Warszawa (1957).

[223] L. D. Landau, I. M. Lifshitz, *Quantum mechanics. Non-relativistic theory*, 3rd edition, Pergamon Press (1991).