Beyond the Gutzwiller Approximation in the Slave-Boson Approach: Inclusion of Fluctuations with the Correct Continuum Limit of the Functional Integral

E. Arrigoni [1] and G. C. Strinati [1,2]
[1]Scuola Normale Superiore, I-56100 Pisa, Italy
[2]Dipartimento di Fisica, Università di Roma "La Sapienza," I-00185 Roma, Italy
(Received 28 June 1993)

We study the single-band Hubbard model via the functional-integral four-slave-boson formulation introduced by Kotliar and Ruckenstein, and show that a proper treatment of the *continuum limit* of the functional integral removes the inconsistencies due to an incorrect handling of this limit so far performed in the literature. It follows from our analysis that the Kotliar and Ruckenstein mean-field solution has to be abandoned in order to include fluctuation corrections in a systematic way. We show explicitly how to construct a suitable alternative solution which admits the inclusion of fluctuations.

PACS numbers: 71.30.+h, 05.30.−d, 71.10.+x

The interest in the Hubbard model(s) to represent strongly correlated fermionic systems has led to the introduction of slave bosons as a convenient means to treat the local (on-site) constraint when the Hubbard repulsion U is large [1]. To implement the local constraint, the slave-boson method has been formulated through a functional-integral approach based on a mixed (fermion-boson) coherent-state representation. In particular, Kotliar and Ruckenstein (KR) have shown that a four-slave-boson representation can be introduced for *any* U [2], yielding at the mean-field level the Gutzwiller results for a half-filled paramagnetic band. The KR method has since been adopted to approach several problems, both at its mean-field level [3] and with the inclusion of fluctuation corrections [4], with the expectation that finite-U models could provide a more flexible description of correlated electronic systems (e.g., the high-T_c materials) than a $U=\infty$ model. This hope has been supported by the finding that the KR mean-field solution gives remarkable agreement with more elaborate (e.g., Monte Carlo [5]) calculations.

With these premises, one would expect the KR mean-field solution to be a good starting point to include fluctuation corrections, the latter being needed anyhow to check the stability of the mean-field results as well as to calculate dynamic quantities defined only beyond the mean-field level. Recently, however, Jolicoeur and Le Guillou [6] have signaled the occurrence of inconsistencies when considering fluctuation corrections within the KR method, although not having identified the actual source of these inconsistencies. We shall, in fact, argue that implementation of the KR method beyond the mean-field level has so far relied on an incorrect procedure for taking the continuum (imaginary time) limit of the functional integral.

The fact that special care must be paid when taking the continuum imaginary time limit within a coherent-state functional-integral approach *in the presence of a slave-boson condensate*, has recently been demonstrated in the context of the $U=\infty$ one-slave-boson problem [7]. In the present paper we concentrate on the implications of taking the correct continuum limit for the four-slave-boson KR method, which prove to be nontrivial since this method presents special peculiarities of its own. It will turn out that the correct continuum limit prevents one from selecting the slave-boson hopping factor z in the form proposed by KR (see below), since fluctuation corrections become ill defined with this particular choice. We shall show how to remedy this shortcoming by proposing a different form of z which is consistent with the correct continuum limit of the functional integral.

For the sake of definiteness, we consider a single-band Hubbard Hamiltonian

$$H=t\sum_{i,\Delta,\sigma}\tilde{f}^{\dagger}_{i,\sigma}\tilde{f}_{i+\Delta,\sigma}+U\sum_{i}\tilde{f}^{\dagger}_{i,\uparrow}\tilde{f}_{i,\uparrow}\tilde{f}^{\dagger}_{i,\downarrow}\tilde{f}_{i,\downarrow}\,, \tag{1}$$

where Δ runs over the "star" of nearest neighbors to site i and σ is a spin label. For finite U, Kotliar and Ruckenstein [2] have mapped the physical fermion operators $\tilde{f}_{i,\sigma}$ in (1) onto the product $f_{i,\sigma}z_{i,\sigma}$ of a (pseudo) fermion $f_{i,\sigma}$ and a bosonic hopping operator $z_{i,\sigma}$. In its *simplest* version $z_{i,\sigma}$ has the form

$$z_{i,\sigma}=s^{\dagger}_{i,-\sigma}d_i+e^{\dagger}_{i}s_{i,\sigma}\,, \tag{2}$$

where e_i, $s_{i,\sigma}$, and d_i refer to empty, singly, and doubly occupied states at site i. This enlargement of the original Fock space requires one to introduce the following constraints:

$$d^{\dagger}_{i}d_i+\sum_{\sigma}s^{\dagger}_{i,\sigma}s_{i,\sigma}+e^{\dagger}_{i}e_i=1\,, \tag{3a}$$

$$f^{\dagger}_{i,\sigma}f_{i,\sigma}-s^{\dagger}_{i,\sigma}s_{i,\sigma}-d^{\dagger}_{i}d_i=0\,. \tag{3b}$$

Fundamental to the KR approach has been the further observation that the simple form (2) for $z_{i,\sigma}$ can be generalized by inserting within *each* pair of operators, for instance, the nonlinear operator

$$R_{i,\sigma}=(1-d^{\dagger}_{i}d_i-s^{\dagger}_{i,\sigma}s_{i,\sigma})^{-1/2}(1-e^{\dagger}_{i}e_i-s^{\dagger}_{i,-\sigma}s_{i,-\sigma})^{-1/2} \tag{4}$$

0031-9007/93/71(19)/3178(4)$06.00

since $R_{i,\sigma}$ acts in this case as the unit operator. This expedient has enabled Kotliar and Ruckenstein to recover in a straightforward way the Gutzwiller solution for a half-filled paramagnetic band, by considering simply the mean-field solution of the corresponding functional integral. Other forms of $R_{i,\sigma}$ different from (4) are, nonetheless, possible in principle.

Enforcement of the constraints can be readily achieved via functional integral techniques by associating the constraints (3) to Lagrange multipliers $\lambda_i^{\rm I}$ and $\lambda_{i,\sigma}^{\rm II}$, in the order. Upon breaking up the (imaginary time) interval $(0,\beta)$ into M steps (where β is the inverse temperature and $M \to \infty$ eventually), one arrives at the following expression for the grand-canonical partition function:

$$\mathcal{Z} = \lim_{M\to\infty} \int \prod_i d\lambda_i^{\rm I} \left[\prod_\sigma d\lambda_{i,\sigma}^{\rm II}\right] \prod_{m=0}^{M-1} d^2 e_{i,m} d^2 d_{i,m} \left[\prod_\sigma d^2 s_{i,\sigma,m} d\bar{f}_{i,\sigma,m}\, df_{i,\sigma,m}\right] \times \exp\left\{-\delta \sum_{m=0}^{M-1} \sum_i (W_{i,m}^{(F)} + W_{i,m}^{(B)} + W_{i,m}^{(FB)} - \lambda_i^{\rm I})\right\}, \qquad (5)$$

where m labels the imaginary time steps, (e,s_σ,d) are complex boson fields, f and $\bar{f}$ are Grassmann variables, and $\delta=\beta/M$ is the elementary time interval. $W_{i,m}^{(F)}$ and $W_{i,m}^{(B)}$ in the action of (5) are functions of the Grassmann variables and of the boson fields, respectively. The mixed fermion-boson term $W_{i,m}^{(FB)}$ derives from the kinetic term in the Hamiltonian (1) and depends on the choice of the bosonic operator $z_{i,\sigma}$:

$$W_{i,m}^{(FB)} = t\sum_{\Delta,\sigma} \bar{f}_{i+\Delta,\sigma,m} z_{i+\Delta,\sigma}^*(m-1,m) \times z_{i,\sigma}(m,m-1) f_{i,\sigma,m-1} \qquad (6)$$

with the notation

$$z_{i,\sigma}(m,m') = e_{i,m}^* R_{i,\sigma}(m,m') s_{i,\sigma,m'} + s_{i,-\sigma,m}^* R_{i,\sigma}(m,m') d_{i,m'}\,. \qquad (7)$$

In particular, with the choice (4) one gets [8]

$$R_{i,\sigma}(m,m') = (1 - d_{i,m}^* d_{i,m'} - s_{i,\sigma,m}^* s_{i,\sigma,m'})^{-1/2} \times (1 - e_{i,m}^* e_{i,m'} - s_{i,-\sigma,m}^* s_{i,-\sigma,m'})^{-1/2}\,. \qquad (8)$$

Although the functional integral representation of the partition function requires one to keep the discretized imaginary time mesh *until the end of the calculation* [9], it has been common practice in the literature of slave bosons to consider the continuum ($\delta \to 0$) limit of the action in (5) at the outset, thereby transferring the $M\to\infty$ limit in (5) under the integral sign [1,2,4,6]. This procedure eliminates the distinction between the two time labels m and m' in (7) and (8). In this way, however, the function (8) unavoidably becomes nonanalytic in certain regions of the integration domain in (5), since the constraint (3a) is enforced only through an average over the interval $(0,\beta)$. Although one may regard this as a merely formal problem, in practice it can lead to unphysical results already at the Gaussian level, as will be explicitly shown below.

It has been argued, however, in the context of the $U=\infty$ one-slave-boson problem that taking the continuum limit of the action at the outset is *pathologic* owing to the presence of the slave-boson condensate [7]. For this reason, the $M\to\infty$ limit has to be properly taken only at the end of the calculation. Otherwise, *wrong* results for the free energy (and for all quantities derived from it) are obtained when fluctuation corrections beyond the saddle point of the functional integral are included. The origin of these problems stems from the commutators of slave-boson operators not being properly accounted for when taking the continuum limit at the outset: In this case, the collapse of pairs of adjacent discretized time labels in the action conspires with the presence of the slave-boson condensate and misses $\frac{1}{2}$ of each commutator.

Concerning the KR choice (8), it is apparent that keeping the two time labels m and m' distinct leads to a further nonanalytic behavior of the functional integral, because in this case the constraint is no longer expressed as a sum of real positive numbers. By this remark, the very discretization of the functional integral prohibits in principle the use of the KR form (8). In practice, we shall show that the KR choice (8) leads to unrealistic results at the Gaussian level, and that these results get consistently worsened by keeping the two time labels distinct.

Preserving correctly the discretized form (5) of the action until the end of the calculation yields (*in addition* to the standard result obtained by taking the continuum limit at the outset [4]) the following contributions to the free energy at the Gaussian level [10,11]:

$$\frac{1}{\mathcal{N}} \Delta F_D = 2\lambda_0^{\rm II} - 2\lambda_0^{\rm I} - \frac{1}{2}U - \frac{1}{4\mathcal{N}} \sum_{\alpha=1}^{4} \frac{\partial F_0}{\partial b_{0\alpha}^2} + \mathcal{F}\,. \qquad (9)$$

In this expression, $\mathcal{N}$ is the number of sites in the system, $b_{0\alpha}$ stand for the four-slave-boson condensate values, $(\lambda_0^{\rm I},\lambda_0^{\rm II})$ for the associated values of the Lagrange multipliers, F_0 for the corresponding free energy, and $\mathcal{F}$ is in general a complicated function of $b_{0\alpha}$ that can be expressed as a linear combination of the first and second derivatives of $R_{i,\sigma}(m,m')$ with respect to $b_{0\alpha}$. In particular, $\mathcal{F}=0$ for the simplest choice $R_{i,\sigma}(m,m')=1$. The general expression for $\mathcal{F}$ will be discussed in detail elsewhere [12].

To show how the above formal considerations affect in practice the results for a model system, we will perform numerical calculations for a one-level two-site model [13], which avoids the unnecessary complications due to the spatial structure and yet keeps those due to the imaginary time discretization. (We expect that the effects described below hold more generally for a multisite system since the spatial structure does not affect the time continuum limit.) The restriction to a two-site model will further enable us to compare our numerical results, obtained within alternative approximations to the functional integral, with the available exact solution.

In Fig. 1 we compare the (canonical) free energy F_c/t for the one-level two-site model (in the zero temperature limit) versus U/t at half filling, as obtained from the exact solution and from alternative approximations to the four-slave-boson method with the KR choice (8) [14]. This comparison shows that the KR mean-field solution (KR) is indeed in rather good agreement with the exact solution (EX), whereas the fluctuation results obtained both by the incorrect (CFL) and by the correct (DFL) time limiting procedure in the functional integral [i.e., without and with the inclusion of the terms (9), respectively] considerably worsen this agreement [15]. This numerical finding confirms the conclusion drawn before by means of formal arguments: In spite of its successes at the mean-field level [2,3,5], the KR form (4) for $R_{i,\sigma}$ is not suitable for a correct mapping of the partition function onto a functional-integral representation.

This conclusion calls for an expression of the bosonic operator $z_{i,\sigma}$ alternative to the KR choice. A possible form for the operator $R_{i,\sigma}$ to be inserted in (2), which satisfies the formal analytic requirements for the functional integral discussed previously and at the same time *reproduces the KR (mean-field) results at half filling*, is obtained by a suitable "linearization" of the form (4):

$$R_{i,\sigma}=[1+x(d_i^\dagger d_i+s_{i,\sigma}^\dagger s_{i,\sigma})][1+x(e_i^\dagger e_i+s_{i,-\sigma}^\dagger s_{i,-\sigma})] \tag{10}$$

with $x=x_0=2(\sqrt{2}-1)\simeq 0.828$. The corresponding results are shown in Fig. 2(a) [14]. Note that the DFL results are now quite close to the exact solution for $0\leq U/t\lesssim 8$, while the agreement is not as good for U larger than the critical value $U_c(\simeq 16t)$ for the Mott-Hubbard transition to occur in the mean-field solution (MF). (The presence of a "cusp" about U_c in the DFL results is of no special concern since it is a characteristic feature of the Gaussian corrections near the transition point.)

The value $x=x_0$ in (10) has been selected to yield $\langle z_{i+\Delta,\sigma}^* z_{i,\sigma}\rangle=1$ at the mean-field level with $U=0$, when the exact solution is trivially known (also for an infinite system). In this context, it is quite remarkable that the fluctuation corrections do not change appreciably the mean-field results with $U=0$ *provided* the terms (9) are included. There is *a priori* no reason, however, to expect the choice (10) with $x=x_0$ to give good results also for large U. In principle, a better agreement could be achieved allowing the parameter x in (10) to be U dependent. For an infinite system this freedom for x can be exploited only when an exact solution or reliable approximations are known (as in the two limits $U=0$ and $U\approx\infty$). In particular, for our model system we find $x=x_\infty\simeq 0.545$ when $U\approx\infty$. Figure 2(b) shows the results obtained with $x=x_0$ for $U\lesssim U_c$ and with $x=x_\infty$ for $U\gtrsim U_c$, plus a smoothing near U_c. Note that the exact solution is rather well approximated by our DFL results

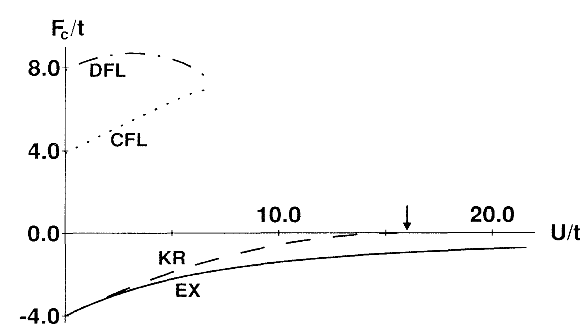

FIG. 1. Free energy (in units of t) for the one-level two-site model versus U/t at half filling, obtained by the exact solution and with the KR choice (4), as explained in the text. The arrow locates the critical value of U/t corresponding to the Mott-Hubbard transition for the mean-field solution.

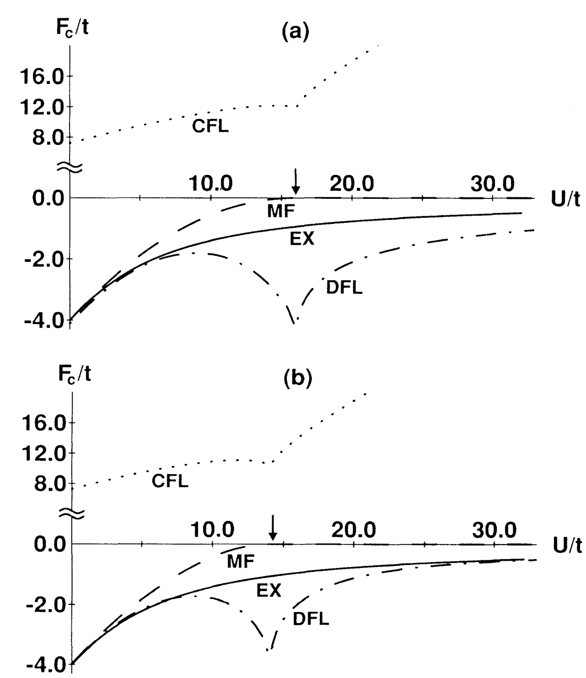

FIG. 2. Free energy (in units of t) for the one-level two-site model versus U/t at half filling, obtained by the exact solution and with the choice (10): (a) $x=x_0$ and (b) $x=P(U/t)x_0+[1-P(U/t)]x_\infty$, where $P(U/t)=[1+\exp(U-U_0)/W]^{-1}$ with $U_0/t=16.0$ and $W/t=4.0$. W/t has been taken of the order of the half-width of the "cusp" in (a). Conventions are as in Fig. 1.

also far from the extremes ($U=0$ and $U\approx\infty$) where x_0 and x_∞ have been selected. We also mention that the "cusp" at U_c in Fig. 2 gets quickly washed out away from half filling, and that the choices x_0 and x_∞ made at half filling provide quite good results even at finite doping [16].

Finally, we remark that, although the results presented in this paper have been restricted to the (zero-temperature) free energy, the correct time limiting procedure and the choice of the bosonic hopping operator should also affect other physical quantities, such as the electronic correlation functions and finite-temperature quantities (i.e., the coefficient of the $T^3 \ln T$ correction to the specific heat considered in Ref. [17]). A complete study of the implications of our procedure for these and related quantities will constitute a significant extension of the present work. Moreover, we expect our procedure to be readily extended to more complicated models, although the choice (10) has been tested here for a simple one-level two-site model. Work along these lines is in progress.

In conclusion, we have shown that unphysical results which are obtained when including fluctuation corrections within the KR four-slave-boson formalism are removed by reconsidering the functional integral formulation in a rigorous fashion with the correct imaginary time limiting procedure. This procedure has forced us to abandon the KR choice for the bosonic hopping operator. Nonetheless, we have been able to preserve the appealing features of the KR mean-field solution through a suitable alternative choice for the bosonic hopping operator, while avoiding the problems encountered with the KR choice beyond the mean-field level.

We thank C. Castellani and R. Raimondi for discussions. One of us (E.A.) gratefully acknowledges partial research support from Europa Metalli-LMI s.p.a.

[1] N. E. Bickers, Rev. Mod. Phys. **59**, 845 (1987), and references quoted therein.

[2] G. Kotliar and A. E. Ruckenstein, Phys. Rev. Lett. **57**, 1362 (1986).

[3] C. Balseiro *et al.*, Phys. Rev. Lett. **62**, 2624 (1989); E. Arrigoni, G. Strinati, and C. Castellani, Phys. Rev. B **41**, 4838 (1990).

[4] M. Lavagna, Phys. Rev. B **41**, 142 (1990); A. Auerbach and K. Levin, Phys. Rev. Lett. **57**, 877 (1986). For a spin-rotation-invariant formulation see T. Li, P. Wölfle, and P. J. Hirschfeld, Phys. Rev. B **40**, 6817 (1989).

[5] L. Lilly, A. Muramatsu, and W. Hanke, Phys. Rev. Lett. **65**, 1379 (1990).

[6] Th. Jolicoeur and J. C. Le Guillou, Phys. Rev. B **44**, 2403 (1991).

[7] E. Arrigoni, C. Castellani, M. Grilli, R. Raimondi, and G. C. Strinati (unpublished).

[8] The square-root operators in (4) get reconciled with the coherent-state representation of the functional integral by taking the normal order of the operator $R_{i,\sigma}$.

[9] Cf., e.g., L. S. Schulman, *Techniques and Applications of Path Integration* (Wiley, New York, 1981).

[10] We work with the *radial gauge* whereby each complex boson field is represented via its amplitude and phase. The procedure commonly adopted (Ref. [1]) absorbs the bosonic phases into a redefinition of the Lagrange multipliers λ (which thus become time dependent) and of the fermion fields. It has been pointed out in Ref. [6] that *only* three boson phases can be absorbed by the three Lagrange multipliers λ^{I} and $\lambda_\sigma^{\mathrm{II}}$ while the phase of one of the four bosons has to be explicitly kept in the functional integral. By convention, we retain the phase of the double occupancy boson d.

[11] The $1/N$ expansion technique, which has been extensively used for the $U=\infty$ one-slave-boson problem, can be suitably generalized to the four-slave-boson problem by retaining only two single occupancy bosons $s_{(+)}$ and $s_{(-)}$ where $\pm$ refer to the sign of the fermion spin projection for half-integer spin $(N-1)/2$. For details see Ref. [12].

[12] E. Arrigoni and G. C. Strinati (unpublished).

[13] A similar model has been recently considered by L. Zhang, J. K. Jain, and V. J. Emery, Phys. Rev. B **40**, 5599 (1992).

[14] Except for the fluctuation results obtained by the incorrect limiting procedure (named CFL in the figures), the free energy properly vanishes in the zero occupancy limit *regardless* of U (with the further exception of the KR choice for $z_{i,\sigma}$ which leads also to divergent DFL results in this limit). This limit thus provides us with a common *reference level* to compare different curves. [The CFL results for F_c/t are instead proportional to U in the zero occupancy limit. This remark also points out the need to consider the additional contributions (9) to the free energy.]

[15] We have verified that this conclusion remains valid also for an infinite system. Note also that the CFL and DFL curves of Fig. 1 terminate at a critical value of $U(\simeq 6.8t)$ where an antiferromagnetic instability develops at the mean-field level.

[16] At finite doping, the CFL results can yield the wrong curvature of F_c/t versus U/t, producing an unphysical *increase* of the (average) number of doubly occupied sites for increasing U/t. This shortcoming always gets remedied by the DFL results, thus giving additional support to the need of including the terms (9) in the free energy.

[17] J. Rasul and T. Li, J. Phys. C **21**, 5119 (1988).