THE GROUND STATE OF A GENERAL ELECTRON-PHONON HAMILTONIAN IS A SPIN SINGLET

J. K. Freericks[(1),†] and Elliott H. Lieb[(2)]

[(1)] *Department of Physics, University of California, Davis, CA 95616*

[(2)] *Departments of Mathematics and Physics, Princeton University, P.O. Box 708, Princeton, NJ, 08544-0708*

Abstract: The many-body ground state of a very general class of electron-phonon Hamiltonians is proven to contain a spin singlet (for an even number of electrons on a finite lattice). The phonons interact with the electronic system in two different ways—there is an interaction with the local electronic charge and there is a functional dependence of the electronic hopping Hamiltonian on the phonon coordinates. The phonon potential energy may include anharmonic terms, and the electron-phonon couplings and the hopping matrix elements may be nonlinear functions of the phonon coordinates. If the hopping Hamiltonian is assumed to have no phonon coordinate dependence, then the ground state is also shown to be unique, implying that there are no ground-state level crossings, and that the ground-state energy is an analytic function of the parameters in the Hamiltonian. In particular, in a finite system any self-trapping transition is a smooth crossover not accompanied by a nonanalytical change in the ground state. The spin-singlet theorem applies to the Su-Schrieffer-Heeger model and both the spin-singlet and uniqueness theorems apply to the Holstein and attractive Hubbard models as special cases. These results hold in all dimensions — even on a general graph without periodic lattice structure.

† Address after Sept. 1, 1994: Department of Physics, Georgetown University, Washington, DC 20057-0995

I. INTRODUCTION

Electrons have a tendency to pair when the effective electron-electron interaction has an attractive region; in particular this occurs when electrons interact by exchanging bosons. The resulting ground state then often sustains either superconducting or charge-density-wave order. The simplest interacting Hamiltonian of this type is one in which electrons interact indirectly with each other via phonons. Migdal[1] analyzed the electron-phonon interaction in the normal state and discovered that in the limit in which the phonon frequency Ω is much smaller than the Fermi energy E_f, the full many-body theory can be described by a first-order, self-consistent Hartree-Fock theory, and that the neglected higher-order diagrams (vertex corrections) usually contribute to order Ω/E_f. This result has been named Migdal's theorem and it classifies those nonadiabatic processes that are typically important for describing low-frequency electron-phonon interactions. Soon thereafter, Eliashberg[2] generalized Migdal's result to the superconducting phase and discovered that a similar first-order self-consistent Hartree-Fock theory would describe superconductivity. Rowell and McMillan[3] subsequently demonstrated that one could directly measure the electron-phonon spectral density from tunneling experiments and then use the formalism of Migdal and Eliashberg to describe all of the remaining properties of the superconducting state. Migdal-Eliashberg theory has been successful in predicting transition temperatures (and other materials properties) of most low temperature superconductors[4,5].

The electron-phonon Hamiltonian considered here is

$$H = \sum_{\sigma} \sum_{x,y \in \Lambda} t_{xy}(\mathbf{q}) c^{\dagger}_{x\sigma} c_{y\sigma} + \sum_{x \in \Lambda} G_x(\mathbf{q})(n_{x\uparrow} + n_{x\downarrow}) + \frac{1}{2} \sum_{j=1}^{\nu} (M_j \Omega_j^2 q_j^2 + \frac{1}{M_j} p_j^2) + V_{\mathrm{an}}(\mathbf{q}). \tag{1.1}$$

We can also add to this an attractive Hubbard type interaction — as discussed later in Sect. III. Our notation is the following: The electrons occupy positions on a finite "lattice" or "graph" Λ, which is some collection of $|\Lambda|$ sites; we emphasize that no specific periodicity or dimensionality is assumed. The operator $c^{\dagger}_{x\sigma}$ ($c_{x\sigma}$) is a creation (annihilation) operator for an electron at lattice site x with z-component of spin $\sigma = \uparrow$ or $\downarrow$. These operators satisfy the anticommutation relations $c^{\dagger}_{x\sigma} c_{y\sigma} + c_{y\sigma} c^{\dagger}_{x\sigma} = \delta_{xy}$ and $c_{x\sigma} c_{y\sigma} + c_{y\sigma} c_{x\sigma} = 0$ for each $\sigma = \uparrow$ or $\downarrow$. It is customary to assume that the up-spin operators also anticommute with the down-spin operators, but it is more convenient for us to assume that they commute, i.e., $c_{x\uparrow}c_{y\downarrow} - c_{y\downarrow}c_{x\uparrow} = 0$, etc.. This change is innocuous [as long as particle number is conserved — which it is with the Hamiltonian (1.1)] and is effected by replacing $c_{x\uparrow}$ by the operator $\exp[i\pi N_\downarrow]c_{x\uparrow}$ with $N_\sigma := \Sigma_{x\in\Lambda} n_{x\sigma}$ and $n_{x\sigma} := c^\dagger_{x\sigma}c_{x\sigma}$; the operator $c^\dagger_{x\uparrow}$ is replaced similarly, but $c_{x\downarrow}$ and $c^\dagger_{x\downarrow}$ are unchanged.

The phonon modes are indexed by j and, for technical simplicity, we assume there are finitely many of them, namely ν. In some special models, such as the Holstein model, defined in (1.9), there is an association between the phonon modes and the lattice sites, but such an association is neither required nor assumed. The phonon coordinates are $q_1, \ldots, q_\nu$ and the momenta are $p_1, \ldots, p_\nu$, denoted collectively by $\mathbf{q}$ and $\mathbf{p}$.

The p_j's and q_j's satisfy the usual commutation relations $[q_j, p_k] = i\delta_{jk}$ and we shall represent these in the usual way as operators on $L^2(\mathbf{R}^\nu)$, the set of square integrable functions of ν variables, by $p_j = -i\mathrm{d}/\mathrm{d}q_j$ (with $\hbar = 1$).

The most general positive-definite quadratic form can always be put in the normal mode form, shown in (1.1), in which the numbers $M_j > 0$ and $\Omega_j > 0$ are, respectively, the masses and frequencies of the corresponding phonon normal modes. For convenience, we explicitly exclude zero frequency modes, which, physically, correspond to center of mass translation. The additional potential term $V_{an}(\mathbf{q})$ includes all non-quadratic terms; it is completely arbitrary except for the assumption that it is bounded below, i.e., $V_{\mathrm{an}}(\mathbf{q}) \geq C$ for some number C, and that $\sum_{j=1}^{\nu} M_j\Omega_j^2 q_j^2 + V_{\mathrm{an}}(\mathbf{q})$ goes to infinity faster than linearly in all directions.

The real hopping matrix $t(\mathbf{q})$, whose elements are $t_{xy}(\mathbf{q})$, is allowed to be an arbitrary measurable function of the phonon coordinates (but *not* the momenta). An important assumption is that $t(\mathbf{q})$ is real and symmetric for each $\mathbf{q}$, i.e., $t_{xy}(\mathbf{q}) = t_{yx}(\mathbf{q})$. We also assume, for convenience, that $\mathrm{Tr}\,|t(\mathbf{q})|$ is finite for every $\mathbf{q}$. [Here, $|t(q)| = \sqrt{t(\mathbf{q})^2}$.] We do not make any assumption about the relative signs and magnitudes of the hopping matrix elements. The reality assumption generically precludes the interaction of the electronic orbital motion with magnetic fields.

The electron-phonon coupling $G_x(\mathbf{q})$ is also an arbitrary real function of the phonon coordinates that couples the phonons to the total electronic charge at lattice site x. For Theorem 1 (existence of singlet ground states) the only assumption about these couplings is lower boundedness of the *total* phonon potential energy, i.e., we assume that the function of $\mathbf{q}$ given by

$$-\mathrm{Tr}\,|t(\mathbf{q})| - 2\sum_{x\in\Lambda} |G_x(\mathbf{q})| + \frac{1}{2}\sum_{j=1}^{\nu} M_j\Omega_j^2 q_j^2 + V_{\mathrm{an}}(\mathbf{q})$$

is bounded below. Usually, one assumes that $G_x(\mathbf{q})$ is a linear function of $\mathbf{q}$, but we do not do so.

The total spin is a conserved quantity of the Hamiltonian H in Eq. (1.1). The spin operators are defined to be the quadratic operators

$$S^z := \frac{1}{2}\sum_{x\in\Lambda}(n_{x\uparrow} - n_{x\downarrow}), \quad S^+ := (S^-)^\dagger := \sum_{x\in\Lambda} c^\dagger_{x\uparrow} c_{x\downarrow}. \tag{1.3}$$

They all commute with the Hamiltonian H. The spin operators satisfy an SU(2) algebra, and the total-spin operator is defined to be the corresponding quadratic Casimir operator $(S_{\mathrm{op}})^2 := (S^z)^2 + \frac{1}{2}S^+S^- + \frac{1}{2}S^-S^+$, with eigenvalues $S(S+1)$. In particular, we are interested in the eigenvalues of the total-spin operator for the ground states of H of the electron-phonon model described in Eq. (1).

Our main result asserts that the ground state of the model in (1.1) has an $S = 0$ ground state, and that the ground state is often unique.

THEOREM 1 (Existence of singlet ground states). *Assume the previously stated conditions on the Hamiltonian H in (1.1). Assume additionally, that the total number of electrons, $2N$, is even. Then among all of the ground states of H there is at least one ground state with total spin $S = 0$.*

For Theorem 2 (uniqueness of the ground state) additional assumptions are needed.

(i) The hopping matrix elements t_{xy} are independent of $\mathbf{q}$. We also assume that Λ is **connected**, i.e., for each x and y in Λ there are sites $x = x_0, x_1, x_2, \ldots, x_n = y$ such that $t_{x_i x_{i+1}} \neq 0$ for all $0 \le i \le n-1$. A **bond** is said to exist between two sites x and y in Λ if $x \neq y$ and if $t_{xy} \neq 0$.

(ii) All the functions of $\mathbf{q}$ appearing in H, i.e., $G_x(\mathbf{q}), V_{\mathrm{an}}(\mathbf{q})$, are differentiable. [Actually, it suffices for them to be locally Hölder continuous with densely defined derivatives.]

(iii) The $G_x(\mathbf{q})$'s are **independent**. By this we mean that for each point $\mathbf{q} \in \mathbf{R}^\nu$ the ν simultaneous equations

$$\sum_{x\in\Lambda} \frac{\partial G_x(\mathbf{q})}{\partial q_j} A_x = 0, \qquad j = 1, 2, \ldots, \nu \tag{1.2}$$

have no common solution other than $A_x = 0$ for all $x \in \Lambda$. In other words the $(|\Lambda| \times \nu)$ matrix $\partial G_x/\partial q_j$ has rank $|\Lambda|$ for each $\mathbf{q} \in \mathbf{R}^\nu$. (Again, it suffices for this to hold only on a dense subset of $\mathbf{R}^\nu$.)

(iv) Every mass, M_j, is finite.

These conditions hold in many models, e.g., the Holstein model[6], but conditions (i) and (iii) do not hold in the Su-Schrieffer-Heeger (SSH) model[7].

THEOREM 2 (Uniqueness of the ground state). *If, (i)-(iv) above are satisfied then the ground state is unique (and hence a nondegenerate spin singlet).*

Remarks.—(1) Theorem 1 has long been conjectured and is consistent with the intuition that the exchange of a boson leads to electron-electron pairing.

(2) The uniqueness theorem establishes that the many-body ground state of H does not have any level crossings for a finite system, thereby establishing the result that the self-trapping transition from a collection of extended polarons to a collection of localized polarons is a smooth crossover, rather than a transition by breaking of analyticity in any finite system.

The proof of these theorems is based upon spin-reflection positivity and is closely related to the analogous proof for the Hubbard model already presented by one of us[8]. A different proof based upon Perron-Frobenius positivity arguments, was given for one-dimensional models[9], but it does not appear to be readily generalizable to the present case.

The SSH model is the special case of (1.1) in which the hopping matrix elements are *linear* functions of the phonon coordinates and the electron-phonon couplings vanish $[G_x(\mathbf{q}) = 0]$. To be more precise consider the original SSH model on a periodic one-dimensional chain[7]

$$H_{SSH} = \sum_\sigma \sum_{i=1}^{|\Lambda|} (t - \delta t[Q_{i+1} - Q_i])(c^\dagger_{i+1\sigma} c_{i\sigma} + c^\dagger_{i\sigma} c_{i+1\sigma}) + \frac{1}{2} \sum_{i=1}^{|\Lambda|} [\kappa(Q_{i+1} - Q_i)^2 + \frac{1}{M} P_i^2] \quad , \tag{1.4}$$

with Q_i (P_i) the local phonon coordinate (momentum) at site i. Transforming to the normal coordinates

$$q_j := \frac{1}{\sqrt{|\Lambda|}} \sum_{k=1}^{|\Lambda|} Q_j \left[\begin{array}{ll} \cos 2\pi k \frac{j - \frac{1}{2}|\Lambda|}{|\Lambda|} & \frac{1}{2}|\Lambda| \le j < |\Lambda| \\ \sin 2\pi k \frac{j - \frac{1}{2}|\Lambda|}{|\Lambda|} & 0 \le j < \frac{1}{2}|\Lambda| \end{array} \right. \tag{1.5}$$

yields the electron-phonon Hamiltonian in the form of Eq. (1.1)

$$H_{SSH} = \sum_{\sigma}\sum_{i=1}^{|\Lambda|}[t - T_{i+1}(\mathbf{q})](c_{i+1\sigma}^{\dagger}c_{i\sigma} + c_{i\sigma}^{\dagger}c_{i+1\sigma}) + \frac{1}{2}\sum_{j=0}^{|\Lambda|-1}[M\Omega_j^2 q_j^2 + \frac{1}{M}p_j^2] \quad , \tag{1.6}$$

with

$$T_i(\mathbf{q}) = \sum_{j=0}^{|\Lambda|/2-1}(-1)^i q_j \sin\frac{2\pi ij}{|\Lambda|} + \sum_{j=|\Lambda|/2}^{|\Lambda|}(-1)^i q_j \cos\frac{2\pi ij}{|\Lambda|} \quad , \tag{1.7}$$

and $\Omega_j^2 = 2\kappa[1+\cos(2\pi j/|\Lambda|)]/M$. Theorem 1 shows that the SSH model always contains a spin-singlet ground state, but the ground state is not necessarily unique. We are aware of no other rigorous results for the SSH model.

The Holstein model[6] is a special case of (1.1) where there is one (internal) normal mode associated with each lattice site (the index j is identical to the index x), the hopping matrix elements have no phonon coordinate dependence, the electron-phonon coupling is linear in the phonon coordinate associated with the lattice site

$$G_x(\mathbf{q}) = g_x q_x \quad , \tag{1.8}$$

and the potential energy is harmonic $[V_{an}(\mathbf{q}) = 0]$. The resulting Holstein Hamiltonian is

$$H_{Hol} = \sum_{\sigma}\sum_{x,y\in\Lambda} t_{xy}c_{x\sigma}^{\dagger}c_{y\sigma} + \sum_{x\in\Lambda} g_x q_x(n_{x\uparrow} + n_{x\downarrow}) + \frac{1}{2}\sum_{x\in\Lambda}(M_x\Omega_x^2 q_x^2 + \frac{1}{M_x}p_x^2) \quad . \tag{1.9}$$

The Holstein model has independent couplings if all g_x are nonzero. The only rigorous result for the Holstein model is that of Löwen [10] for *one* electron. In this case, it has been shown that the ground state is nondegenerate and analytic if the lattice is bipartite.

In the static limit, where all of the phonon masses become infinite, but the spring constant remains finite

$$M_x \to \infty \quad , \quad \kappa_x = M_x\Omega_x^2 = \text{finite} \quad , \tag{1.10}$$

the phonon kinetic energy terms $\sum_x p_x^2/2M_x$ do not contribute to the Hamiltonian (1.9). The up- and down-spin electrons become independent and the Holstein model maps onto a Falicov-Kimball model[11] with spin one-half conduction electrons and a *continuous* static field q_x. Techniques similar to those used in the spinless Falicov-Kimball model[12] may be used to show that the ground state is a commensurate charge-density-wave at half filling[13,14]. The static limit of the Holstein model has also been investigated by other methods[15] and shown to possess insulating bipolaronic charge-density-wave order at large enough coupling. The ground state has also been shown to be nonanalytic in one dimension[15].

In the instantaneous limit, where the phonon frequency and electron-phonon coupling become infinite, but their ratio remains finite,

$$g_x \to \infty \quad , \quad \Omega_x \to \infty \quad , \quad \frac{g_x}{\Omega_x} = \text{finite} \quad , \tag{1.11}$$

the Holstein model maps onto an attractive Hubbard model[16]. This mapping is illustrated by completing the square in Eq. (1.9)

$$H_{Hol} = \sum_\sigma \sum_{x,y\in\Lambda} t_{xy} c^\dagger_{x\sigma} c_{y\sigma} - \frac{1}{2}\sum_{x\in\Lambda} U_x (n_{x\uparrow} + n_{x\downarrow})^2$$

$$+\frac{1}{2}\sum_{x\in\Lambda}(M_x\Omega_x^2[q_x + \frac{g_x}{M_x\Omega_x^2}\{n_{x\uparrow}+n_{x\downarrow}\}]^2 + \frac{1}{M_x}p_x^2) \quad , \tag{1.12}$$

with the electron-electron interaction U_x defined to be

$$U_x := \frac{g_x^2}{M_x\Omega_x^2} \quad . \tag{1.13}$$

In this instantaneous limit the remaining electron and phonon terms in the Hamiltonian decouple because $g_x/M_x\Omega_x^2 \to 0$, and one is left with an attractive Hubbard model

$$H_{Hub} = \sum_\sigma \sum_{x,y\in\Lambda} \bar{t}_{xy} c^\dagger_{x\sigma} c_{y\sigma} - \sum_{x\in\Lambda} U_x n_{x\uparrow} n_{x\downarrow} + \frac{1}{2}\sum_{x\in\Lambda}(M_x\Omega_x^2 q_x^2 + \frac{1}{M_x}p_x^2) \quad , \tag{1.14}$$

with

$$\bar{t}_{xy} := t_{xy} - \frac{1}{2}U_x\delta_{xy} \quad . \tag{1.15}$$

The attractive Hubbard model is already known to have a unique spin-singlet ($S = 0$) ground state for an even number of electrons on a finite lattice[8].

In Section II, the proofs of Theorems 1 and 2 are presented for the Hubbard model in order to clarify the results of Ref. 8 and to define our notation and methodology. Section III contains the proofs of these theorems for the electron-phonon Hamiltonian. A discussion of the results follows in Section IV.

II. ATTRACTIVE HUBBARD MODEL PROOFS

Proofs are presented for the results in Ref. 8 in order to clarify the previous work and to define the notation and current methodology. We begin with a proof of Theorem 1. The attractive Hubbard model Hamiltonian is given by the electronic terms in Eq. (1.14) with each $U_x \geq 0$ and the bar dropped from the hopping matrix (the hopping matrix elements have no phonon coordinate dependence here).

Proof of Theorem 1 for the attractive Hubbard model.—Both the total spin operator S^2 and the z-component of spin S^z are conserved quantities of the Hubbard model (1.14) and restriction can be made to the $S^z = 0$ subspace (without loss of generality), because every eigenstate with total-spin S can be rotated into the subspace with $S^z = 0$ without changing its energy. Therefore, we can assume there are N electrons of each spin (up and down).

It is convenient to use first quantized notation. We denote the coordinates of the up-spin electrons with X which really is an N-tuple $X = (x_1, x_2, \ldots, x_N)$ where each $x_i \in \Lambda$. Similarly Y denotes the coordinates of the down-spin electrons. Any wave function Ψ is a function of both coordinates $\Psi(X,Y)$ and it is antisymmetric in $\{x_i\}_{i=1}^N$ and antisymmetric in $\{y_i\}_{i=1}^N$. Note that if $\Psi(X,Y)$ is an eigenfunction, then so is $\Psi(Y,X)$ and $\Psi(Y,X)^*$. (It is here that the condition that the hopping matrix elements t_{xy} are real is used.) Instead of considering an eigenfunction $\Psi(X,Y)$ it is convenient to consider $\Psi(X,Y)+\Psi(Y,X)^*$ and $i[\Psi(X,Y) - \Psi(Y,X)^*]$. In other words we can, without loss of generality, assume that our eigenfunction, viewed as a matrix indexed by X and Y, is self adjoint (but not necessarily real).

$$\Psi(X,Y) = \Psi(Y,X)^*. \tag{2.1}$$

The dimension, d, of this matrix Ψ is

$$d = \binom{|\Lambda|}{N}. \tag{2.2}$$

Any self-adjoint matrix can be expanded in terms of its eigenfunctions. Thus,

$$\Psi(X,Y) = \sum_{\alpha=1}^{d} w_\alpha \phi_\alpha(X)\phi_\alpha(Y)^* \tag{2.3}$$

where the ϕ_α's are an orthonormal set of functions (but antisymmetric in their argument X), and the w_α's are real numbers. Our aim is to show that the w_α's can all be chosen to be nonnegative. This will conclude the proof because it implies that

$$\sum_X \psi(X,X) > 0, \tag{2.4}$$

which implies that $\psi(X_0, X_0)$ is positive for at least one X_0. This means that the wave function does not vanish when the up-spin electrons and the down-spin electrons are at precisely the same locations — and this is a singlet state. Thus ψ has a nonvanishing component in the $S=0$ sector.

To show that the w_α's can be taken nonnegative, let us write out the energy using the decomposition of Ψ in (2.3). One easily computes

$$\langle\Psi|H|\Psi\rangle = 2\sum_{\alpha=1}^{d} w_\alpha^2 \langle\phi_\alpha|K|\phi_\alpha\rangle - \sum_{x=1}^{|\Lambda|} U_x \sum_{\alpha=1}^{d}\sum_{\beta=1}^{d} w_\alpha w_\beta \big|\langle\phi_\alpha|L_x|\phi_\beta\rangle\big|^2 \tag{2.5}$$

$$\langle\Psi|\Psi\rangle = \sum_{\alpha=1}^{d} w_\alpha^2 \langle\phi_\alpha|\phi_\alpha\rangle \tag{2.6}$$

where the d-dimensional matrices K and L_x are defined as follows: K is the first-quantized version of $\sum_{xy} t_{xy} c_x^\dagger c_y$ (no spin here) and L_x is the first-quantized version of n_x. More explicitly, the matrix elements appearing in (2.5) and (2.6) are constructed in the following manner: The inner product between two arbitrary vectors $\phi_1(X)$ and $\phi_2(X)$ is

$$\langle\phi_2|\phi_1\rangle = \sum_X \phi_2(X)^*\phi_1(X). \tag{2.7}$$

The kinetic energy matrix elements satisfy

$$\langle\phi_2|K|\phi_1\rangle = \sum_{xy} t_{xy} \sum_{j=1}^{N} \sum_X \phi_2(x_1,\ldots,x,\ldots,x_N)^*\phi_1(x_1,\ldots,y,\ldots,x_N) \tag{2.8}$$

where the argument of ϕ_2 agrees with the argument of ϕ_1 everywhere except at the j^{th} index where the site index is x for ϕ_2 and y for ϕ_1. The number operator matrix elements satisfy

$$\langle\phi_2|L_x|\phi_1\rangle = \sum_X \phi_2(X)^*\phi_1(X) \sum_{j=1}^{N} \delta_{x_j,x} \tag{2.9}$$

For the purpose of Theorem 1, the explicit values of these matrix elements are unimportant. The only thing one has to note about (2.5) and (2.6) is that replacing every w_α by $|w_\alpha|$ cannot increase the energy. The first term in (2.5) and the inner product of (2.6) stay the same, while the second term in (2.5) can only improve (if it changes at all). Thus if Ψ, given by (2.3) is a ground state of the Hubbard model then so is $|\Psi|$ which is constructed by replacing $w_\alpha \to |w_\alpha|$ in (2.3). Note that $|\Psi|(X,Y)$ is not equal to $|\Psi(X,Y)|$ in general, but corresponds to $|\Psi| = \sqrt{\Psi^2}$ in the sense of matrices. **Q.E.D.**

In order to prove Theorem 2 we first need to establish a lemma and in order to state the lemma a definition is needed. If $X = (x_1, \dots, x_N)$ we define the operator Π^X to be

$$\Pi^X := L_{x_1} L_{x_2} \dots L_{x_N} \tag{2.10}$$

Since the different L_x's commute, the ordering of factors in (2.10) is unimportant. In second quantized notation $\Pi^X = n_{x_1} \dots n_{x_N}$ which shows that Π^X is an orthogonal projector, i.e., $\Pi^X = (\Pi^X)^\dagger$ and $(\Pi^X)^2 = \Pi^X$. It is also obvious that Π^X is a *one-dimensional* projector. Furthermore, if X and X' differ only by a permutation then $\Pi^X = \Pi^{X'}$. The matrix elements of Π^X [analogous to (2.9)] satisfy

$$\langle \phi_2 | \Pi^X | \phi_1 \rangle = \phi_2(X)^* \phi_1(X) N!. \tag{2.11}$$

[The $N!$ in (2.11) may appear mysterious, but it is not. The reason is that if ϕ_1 and ϕ_2 are normalized vectors concentrated at X and all of its $N!$ permutations and at no other X, then $|\phi_1(X)|^2 = |\phi_2(X)|^2 = 1/N!$. Thus, both sides of Eq.(2.11) are equal in magnitude to 1.]

LEMMA (Connectivity of the single-spin configuration space). *Assume that the lattice Λ is connected as explained above. Then the single-spin configuration space is connected by the kinetic energy matrix K. That is to say, given points $X = (x_1, \dots, x_N), Y = (y_1, \dots, y_N)$ in the single-spin configuration space, there exists a chain of m points $\{Y = X_m, X_{m-1}, \dots, X_2, X_1 = X\}$ in the configuration space, such that the product of matrix elements satisfies*

$$\Pi^{X_m} K \Pi^{X_{m-1}} \dots \Pi^{X_3} K \Pi^{X_2} K \Pi^{X_1} \neq 0, \tag{2.12}$$

with $X_1 \neq X_2 \neq X_3 \neq \dots \neq X_m$.

Proof: We first consider a geometric question. Place N unlabelled markers on the points $(x_1, \ldots, x_N)$ of Λ. The goal is to move these markers, one at a time, across bonds of the lattice Λ to a final set of points $(y_1, \ldots, y_N)$ in such a way that at no step is there ever a doubly occupied site of the lattice. We want to emphasize that it is *not* necessary that the marker that was first at x_1 ends up at y_1, we require *only* that in the final configuration the sites $(y_1, \ldots, y_N)$ are occupied by some marker.

To do this we apply the following algorithm repeatedly — at most N times. Look for the smallest i such that the site y_i is unoccupied. Look for the smallest j such the site x_j is not in the set $\{y_1, \ldots y_N\}$. We will move a marker from the point x_j and establish a marker at the point y_i in such a way that the other occupied sites of the final state are identical to the other occupied sites of the initial state. (Once again, we point out that the marker originally at x_j need not end up at y_i, and the other markers may be moved in this process.) To achieve this we choose a connected path P in the lattice $\Lambda, P = (x_j = z_1, z_2, \ldots z_k = y_i)$ from site x_j to y_i. Such a path exists by hypothesis. If there are no markers on the sites $z_2, \ldots, z_{k-1}$, then we simply move the marker at x_j along the path P to y_i. Suppose on the contrary, that there are some other markers on this path P. Let l be the largest number such that z_l has a marker on it. Then simply move this marker along the path to the site y_i, thereby achieving two things: a marker on y_i and one less marker along the path P. We then move in turn each marker on the path P to the location of the previously moved marker. This completes the description of the algorithm and answers the geometric question.

To prove the lemma itself, we first note that $\langle\phi_2|\Pi^{X_1}K\Pi^{X_2}|\phi_1\rangle = 0$ for all X_1 and X_2 unless there is a permutation of X_1 such that after the permutation there exists some $1 \leq j \leq N$ such that $x_i^1 = x_i^2$ for all $i \neq j$. In this case

$$\langle\phi_2|\Pi^{X_1}K\Pi^{X_2}|\phi_1\rangle = t_{ab}\phi_2(X_1)^*\phi_1(X_2)N! \tag{2.13}$$

where $x_j^1 = a$ and $x_j^2 = b$.

The configurations $X_1, X_2, \ldots X_m$ used in Eq. (2.12) will be the X's determined by the sequence of moves in the geometric discussion above. (Note that although the markers were indistinguishable there we can, if we wish, put numbers on them. In general the final state in this case will not be identically the state $Y = (y_1, \ldots y_N)$ but will be some permutation $\widetilde{Y}$ of the final state. On the other hand $\Pi^Y = \Pi^{\widetilde{Y}}$ because the L_x's commute.) What remains to be shown is that the operator

$$\Pi^{X_m} K \Pi^{X_{m-1}} \cdots \Pi^{X_2} K \Pi^{X_1} = (\Pi^{X_m} K \Pi^{X_{m-1}})(\Pi^{X_{m-1}} K \Pi^{X_{m-2}}) \cdots (\Pi^{X_2} K \Pi^{X_1}) \quad (2.14)$$

is nonvanishing. Let $\phi_1, \ldots, \phi_d$ be the orthonormal basis proportional to antisymmetrized delta-functions in the configuration space. ¿From (2.13) we see that when we expand (2.14) in these intermediate states, that a factor such as $\langle \phi_i | \Pi^{X_k} K \Pi^{X_{k-1}} | \phi_j \rangle$ can be nonzero for only one i and j, and this contribution is equal to t_{ab} (up to an overall sign) where a and b denote the two indices where X_k and X_{k-1} differ. So, in short, there will be exactly one matrix element of $\langle \phi_i | \Pi^{X_m} K \Pi^{X_{m-1}} \ldots \Pi^{X_2} K \Pi^{X_1} | \phi_j \rangle$, and it will be a product of t_{xy}'s all of which are nonzero. **Q.E.D.**

We turn now to the proof of Theorem 2 for the Hubbard model. The four assumptions (i)-(iv) for the electron-phonon model simplify to the two assumptions: (a) Λ is connected; (b) every U_x is positive ($U_x > 0$).

Proof of Theorem 2 for the Hubbard model. Suppose Ψ_1 and Ψ_2 are orthogonal ground states. We can assume both Ψ_1 and Ψ_2 are Hermitian. Then $\Psi = \Psi_1 + \lambda \Psi_2$ is a ground state for all real λ. Moreover Ψ, if viewed as a matrix, cannot be either positive or negative semidefinite for all real λ. Hence for some choice of λ we have that the two ground states $\Psi_\pm := \frac{1}{2}(|\Psi| \pm \Psi)$, are both nonzero, and are both ground states, since $|\Psi|$ is a ground state from Theorem 1. In particular Ψ_+ satisfies the matrix Schrödinger equation

$$K\Psi_+ + \Psi_+ K - \sum_{x=1}^{|\Lambda|} U_x L_x \Psi_+ L_x = e\Psi_+, \quad (2.15)$$

and is a positive semidefinite matrix.

We define $\mathcal{H}_+$ to be the range of the matrix Ψ_+, which is a subspace of $\mathcal{H} = \mathbf{C}^d$. We also define $\mathcal{H}_\perp$ to be the orthogonal component of $\mathcal{H}_+$ and $\Pi_\perp$ to be the projector onto $\mathcal{H}_\perp$. By assumption $\mathcal{H}_+$ and $\mathcal{H}_\perp$ are both nontrivial.

Multiply the Schrödinger equation (2.15) on the left and on the right by $\Pi_\perp$ to yield $\sum_x U_x \Pi_\perp L_x \Psi_+ L_x \Pi_\perp = 0$ since $\Pi_\perp \Psi_+ = \Psi_+ \Pi_\perp = 0$. This implies further that

$$\Psi_+ L_x \Pi_\perp = 0 \quad (2.16)$$

since every U_x is positive and Ψ_+ is a positive semidefinite matrix. Eq. (2.16) implies that the matrix L_x does not connect the subspaces $\mathcal{H}_+$ and $\mathcal{H}_\perp$. Now multiply (2.15) on the right by $\Pi_\perp$ and use (2.16) to discover that

$$\Psi_+ K \Pi_\perp = 0. \tag{2.17}$$

Both (2.16) and (2.17) show that $\mathcal{H}_+$ and $\mathcal{H}_\perp$ are invariant subspaces of the matrices K and L_x.

We will now use the lemma to show that $\mathcal{H}_\perp$ is trivial, thereby establishing a contradiction. The operators Π^X defined in (2.10) are *one-dimensional* projectors in $\mathcal{H}$ and satisfy

$$(N!)^{-1} \sum_X \Pi^X = \mathbf{1}. \tag{2.18}$$

Thus $V^+ = \Pi^X \Psi^+$ is nonzero for some X. Likewise, if $\phi \in \mathcal{H}_\perp$ then $V^\perp = \Pi^Y \phi \neq 0$ for some Y. By the lemma, the rank one operator in (2.12), call it A, is nonzero and satisfies $(V^\perp, AV^+) \neq 0$. This is a contradiction since A, being a product of L_x's and K, has $\mathcal{H}_+$ as an invariant subspace. **Q.E.D.**

III. ELECTRON-PHONON MODEL PROOFS

The proof of Theorem 1 employs the same methods as utilized for the Hubbard model in the preceding section, but now all of the expansion coefficients and basis functions have an implicit $\mathbf{q}$ dependence. The variational principle that shows that the ground state includes a spin-singlet state will arise from the kinetic energy terms for the phonons, since there is no direct electron-electron interaction here (however, see the remark at the end of the proof of Theorem 2 about including *attractive* electron-electron interactions).

Proof of Theorem 1 for the electron-phonon model. Any wavefunction $\Psi(\mathbf{q})$ can be thought of as a matrix-valued function of $\mathbf{q} \in \mathbf{R}^\nu$. By $SU(2)$ invariance of (1.1) we can, as in the preceeding proof, assume that there are equal numbers of up-spin and down-spin electrons, namely N of each kind. The dimension d of the matrix $\Psi(\mathbf{q})$ is then

$$d = \binom{|\Lambda|}{N}. \tag{3.1}$$

As before, by taking transposes and complex conjugates, we can restrict our discussion to the case where $\Psi(\mathbf{q})$ is a Hermitian matrix for all $\mathbf{q}$ (it is here that we use the condition that the hopping matrix elements t_{xy} are real because, without this condition, the complex conjugate of Ψ will generally have a different energy from that of Ψ). Note that $\Psi(\mathbf{q})$ is Hermitian, but it need not be real.

We can write the Schrödinger equation for Ψ in the following generic way

$$-\sum_{j=1}^{\nu} \frac{1}{2M_j} \frac{\partial^2}{\partial q_j^2} \Psi(\mathbf{q}) + V(\mathbf{q})\Psi(\mathbf{q}) + \Psi(\mathbf{q})V(\mathbf{q}) = e\Psi(\mathbf{q}). \tag{3.2}$$

This equation is to be understood in the following sense: $\Psi(\mathbf{q})$ is a matrix-valued function and so are its second derivatives; $V(\mathbf{q})$ is also a matrix-valued function which is self-adjoint for every $\mathbf{q}$; the two terms $V\Psi + \Psi V$ (with matrix multiplication being understood) include all of the terms in $\mathcal{H}$ besides the phonon kinetic energy, $\sum_{j=1}^{\nu} p_j^2/2M_j$; of course e is the energy eigenvalue.

Associated with (3.2) is a variational expression which we can write as $\mathcal{E}(\Psi)/\langle\Psi|\Psi\rangle$. The denominator has the form

$$\langle\Psi|\Psi\rangle = \int_{\mathbf{R}^\nu} \mathrm{Tr}\Psi(\mathbf{q})^2 \mathrm{d}\mathbf{q}. \tag{3.3}$$

The numerator is

$$\mathcal{E}(\Psi) = \int_{\mathbf{R}^\nu} \left\{ \sum_{j=1}^{\nu} \frac{1}{2M_j} \mathrm{Tr}[\partial_j \Psi(\mathbf{q})]^2 + 2\mathrm{Tr}V(\mathbf{q})\Psi(\mathbf{q})^2 \right\} \mathrm{d}\mathbf{q}, \tag{3.4}$$

where ∂_j denotes the partial derivative $\partial/\partial q_j$.

Our strategy is to replace the matrix $\Psi(\mathbf{q})$, for every $\mathbf{q}$, by its absolute value in the matrix sense, i.e.,

$$|\Psi(\mathbf{q})| = \sqrt{\Psi(\mathbf{q})^2}. \tag{3.5}$$

We note that the norm satisfies $\langle\Psi|\Psi\rangle = \langle|\Psi|\,|\,|\Psi|\rangle$, and that the $V\Psi^2$ term is evidently unchanged. In the appendix we prove that this substitution does not increase the integral of $\mathrm{Tr}[\partial_j\Psi(\mathbf{q})]^2$. There are some nontrivial technical issues here caused by the fact that $\partial_j\Psi(\mathbf{q})$ and $\partial_j|\Psi(\mathbf{q})|$ may only be distributional derivatives, but these issues are fully resolved in the appendix.

Since, by definition of the ground-state energy, $\mathcal{E}(\Psi)$ cannot decrease when Ψ is replaced by $|\Psi|$, we conclude that $|\Psi|$ is also a ground state of the Hamiltonian. As in the

Hubbard model proof, we conclude that $|\Psi(\mathbf{q})|$ has a nonzero projection onto the $S=0$ subspace, i.e., if every diagonal element of the matrix $|\Psi(\mathbf{q})|$ vanished for almost every value of $\mathbf{q}$, $|\Psi(\mathbf{q})|$ would be the zero function, which it is not. **Q.E.D.**

Proof of Theorem 2 for the electron-phonon model. Suppose there are two ground states $\Psi_1(\mathbf{q})$ and $\Psi_2(\mathbf{q})$ which are Hermitian and linearly independent. Then $\Psi_1+\lambda\Psi_2$ is a nonzero ground state for every real λ, and as λ is varied from $-\infty$ to $+\infty$ there will be values of λ for which $\Psi=\Psi_1+\lambda\Psi_2$ has both a negative and positive spectrum for a set of $\mathbf{q}$'s of positive measure. Then $|\Psi|$ is also a ground state, as are $\Psi_+ := \frac{1}{2}(|\Psi|+\Psi)$ and $\Psi_- := \frac{1}{2}(|\Psi|-\Psi)$. Note that $\Psi, |\Psi|, \Psi_+$ and Ψ_- are all nonzero functions that satisfy the Schrödinger equation (3.2). Indeed $\Psi=\Psi_+-\Psi_-$. Moreover, both $\Psi_+(\mathbf{q})$ and $\Psi_-(\mathbf{q})$ are positive semidefinite matrices for all values of $\mathbf{q}$. ¿From the fact that the matrix valued function $V(\mathbf{q})$ appearing in (3.2) is differentiable, elliptic regularity theory applied to the Schrödinger equation tells us that $\Psi(\mathbf{q})$ is twice continuously differentiable [actually only Hölder continuity of $V(\mathbf{q})$ suffices for this conclusion].

At each point $\mathbf{q}$, the vector space $\mathbf{C}^d$, on which $\Psi(\mathbf{q})$ operates, is naturally the direct sum of three subspaces (some of which might be empty). These subspaces are denoted by $\mathcal{H}_+(\mathbf{q}), \mathcal{H}_-(\mathbf{q})$, and $\mathcal{H}_0(\mathbf{q})$. $\mathcal{H}_+(\mathbf{q})$ is the spectral subspace of $\Psi_+(\mathbf{q})$, i.e. consists of all linear combinations of the nonzero eigenvectors of $\Psi_+(\mathbf{q})$; $\mathcal{H}_-(\mathbf{q})$ is the spectral subspace of $\Psi_-(\mathbf{q})$; and $\mathcal{H}_0(\mathbf{q})$ is the orthogonal complement of $\mathcal{H}_+(\mathbf{q})\oplus\mathcal{H}_-(\mathbf{q})$, i.e., the subspace of all linear combinations of the zero eigenvectors of $\Psi_+(\mathbf{q})$ and $\Psi_-(\mathbf{q})$. Of course $\mathbf{C}^d=\mathcal{H}_+(\mathbf{q})\oplus\mathcal{H}_0(\mathbf{q})\oplus\mathcal{H}_-(\mathbf{q})$.

Let $d_0(\mathbf{q})$ denote the dimension of $\mathcal{H}_0(\mathbf{q})$ and let d_0 denote the minimum of $\{d_0(\mathbf{q}) : \mathbf{q}\in\mathbf{R}^\nu\}$. Since there are only finitely many values for $d_0(\mathbf{q})$ there is a point $\mathbf{q}_0\in\mathbf{R}^\nu$ for which $d_0(\mathbf{q}_0)=d_0$. By definition the matrix $\Psi(\mathbf{q}_0)$ has d_0 zero eigenvalues, and the positive eigenvalues are separated from zero by a gap Δ. Since Ψ, and hence the eigenvalues of Ψ, are continuous, there is some ball $\mathcal{B}'$ centered at $\mathbf{q}_0$ with radius r such that the dimension of $d_0(\mathbf{q})\le d_0$ for every $\mathbf{q}\in\mathcal{B}'$. Since, however d_0 is the minimum of $d_0(\mathbf{q})$, we conclude that $d_0(\mathbf{q})=d_0$ for every $\mathbf{q}\in\mathcal{B}'$.

Now let us study the contour integral

$$\Pi_\perp(\mathbf{q}) := \frac{1}{2\pi i}\int_\Gamma \frac{1}{\Psi(\mathbf{q})-z}\mathrm{d}z \tag{3.6}$$

where the contour Γ runs from $-\infty$ just below the negative real axis, goes vertically upward when the real part of $z = \Delta/2$ and returns to $-\infty$ just above the negative real axis.

First we observe that $\Pi_\perp(\mathbf{q}_0)$ is the projector onto the orthogonal complement of $\mathcal{H}_+(\mathbf{q}_0)$, namely onto $\mathcal{H}_\perp(\mathbf{q}_0) = \mathcal{H}_0(\mathbf{q}_0) \oplus \mathcal{H}_-(\mathbf{q}_0)$. Furthermore, the zero eigenvalues of $\Psi_+(\mathbf{q})$ do not move from zero as long as $\mathbf{q} \in \mathcal{B}'$, as we have just seen above. Therefore, $\Pi_\perp(\mathbf{q})$ continues to be the projector onto $\mathcal{H}_\perp(\mathbf{q})$ as long as the positive eigenvalues of $\Psi(\mathbf{q})$ are greater than $\Delta/2$, and hence do not intersect the contour Γ. Since Ψ is continuous we conclude there is a smaller ball $\mathcal{B} \subset \mathcal{B}'$ centered at $\mathbf{q}$ in which $\Pi_\perp(\mathbf{q})$ continues to be the projector onto $\mathcal{H}_\perp(\mathbf{q})$. Since $\Psi(\mathbf{q})$ is twice continuously differentiable, it is a trivial matter to show that we can differentiate under the integral sign in (3.6) and conclude that $\Pi_\perp(\mathbf{q})$ *is a twice continuously differentiable matrix-valued function on* $\mathcal{B}$.

Now we compute some derivatives in this ball $\mathcal{B}$. We start with the observation that $\Psi_+(\mathbf{q})\Pi_\perp(\mathbf{q}) = 0$ for all $\mathbf{q}$ in $\mathbf{R}^\nu$. The following identities hold in $\mathcal{B}$: (we suppress the $\mathbf{q}$ dependence):

$$(\partial_j \Psi_+)\Pi_\perp + \Psi_+ \partial_j \Pi_\perp = 0 \tag{3.7}$$

$$(\partial_j^2 \Psi_+)\Pi_\perp + 2(\partial_j \Psi_+)\partial_j \Pi_\perp + \Psi_+ \partial_j^2 \Pi_\perp = 0. \tag{3.8}$$

If the Schrödinger equation (3.2) for $\Psi_+(\mathbf{q})$ is multiplied on the left and on the right by $\Pi_\perp(\mathbf{q})$ we have (since $\Psi_+\Pi_\perp = 0 = \Pi_\perp\Psi_+$)

$$\sum_{j=1}^{\nu} \frac{1}{2M_j} \Pi_\perp (\partial_j^2 \Psi_+) \Pi_\perp = 0, \tag{3.9}$$

or, combining (3.9) with (3.8) we discover that

$$\sum_{j=1}^{\nu} \frac{1}{2M_j} \Pi_\perp (\partial_j \Psi_+) \partial_j \Pi_\perp = 0. \tag{3.10}$$

Now multiply (3.7) on the left by $\partial_j \Pi_\perp / 2M_j$ and sum over j. The first term vanishes because of the adjoint of (3.10). The second yields

$$\sum_{j=1}^{\nu} \frac{1}{2M_j} \partial_j \Pi_\perp \Psi_+ \partial_j \Pi_\perp = 0. \tag{3.11}$$

Since $\Psi_+(\mathbf{q})$ is a positive semidefinite matrix (and all $M_j < \infty$), we conclude from (3.11) that

$$\Psi_+ \partial_j \Pi_\perp = 0, \qquad j = 1, \ldots, \nu \tag{3.12}$$

for all $\mathbf{q} \in \mathcal{B}$. Eq. (3.12) states that the range of the derivatives of $\Pi_\perp$ lies in $\mathcal{H}_\perp$, or

$$\partial_j \Pi_\perp = \Pi_\perp \partial_j \Pi_\perp, \tag{3.13}$$

and therefore that each term in (3.7) separately vanishes, i.e.,

$$(\partial_j \Psi_+)\Pi_\perp = 0. \tag{3.14}$$

Differentiating (3.14) yields

$$(\partial_j^2 \Psi_+)\Pi_\perp + (\partial_j \Psi_+)\partial_j \Pi_\perp = 0. \tag{3.15}$$

But $(\partial_j \Psi_+)\partial_j \Pi_\perp = (\partial_j \Psi_+)\Pi_\perp \partial_j \Pi_\perp = 0$ from (3.13) and (3.14), so we finally conclude that the range of the second derivative of Ψ_+ lies in $\mathcal{H}_+$, or

$$(\partial_j^2 \Psi_+)\Pi_\perp = 0 \tag{3.16}$$

for all $q \in \mathcal{B}$.

If we multiply the Schrödinger equation (3.2) for Ψ_+ on the right by $\Pi_\perp$ and use the identity (3.16) we find

$$\Psi_+(\mathbf{q}) \left[K(\mathbf{q}) + \sum_{x=1}^{|\Lambda|} L_x G_x(\mathbf{q}) \right] \Pi_\perp(\mathbf{q}) = 0. \tag{3.17}$$

This says that the matrix $W(\mathbf{q}) := K(\mathbf{q}) + \sum_{x=1}^{|\Lambda|} L_x G_x(\mathbf{q})$ does not connect $\mathcal{H}_+$ with $\mathcal{H}_\perp$. Differentiating (3.17) shows that

$$(\partial_j \Psi_+) W \Pi_\perp + \Psi_+ (\partial_j W) \Pi_\perp + \Psi_+ W \partial_j \Pi_\perp = 0. \tag{3.18}$$

But we know that the range of the derivatives of Ψ_+ lies in $\mathcal{H}_+$ and that of the derivatives of $\Pi_\perp$ lies in $\mathcal{H}_\perp$, so the first and last terms of (3.18) vanish because W does not connect $\mathcal{H}_+$ with $\mathcal{H}_\perp$. Since the hopping matrix elements do not have any $\mathbf{q}$ dependence by assumption and the derivative of the electron-phonon couplings is a rank $|\Lambda|$ matrix by assumption, we find

$$\Psi_+ L_x \Pi_\perp = 0 \qquad x = 1, \dots, |\Lambda| \tag{3.19}$$

and, from (3.17),

$$\Psi_+ K \Pi_\perp = 0. \tag{3.20}$$

These two identities imply that both $\mathcal{H}_+(\mathbf{q})$ and $\mathcal{H}_\perp(\mathbf{q})$ are invariant subspaces of the matrices K and L_x for all $\mathbf{q} \in \mathcal{B}$. Exactly as in the Hubbard model proof, we conclude that for every $\mathbf{q} \in \mathcal{B}$ one of the two alternatives holds: either $\mathcal{H}_+(\mathbf{q}) = \{0\}$ or $\mathcal{H}_-(\mathbf{q}) = \{0\}$. Since the functions Ψ_+ and Ψ_- are continuous, the set on which Ψ_+ is nonzero is open and the set on which Ψ_- is nonzero is open. Therefore $\mathcal{B}$ contains an open set in which either $\Psi_+(\mathbf{q})$ is identically zero or $\Psi_-(\mathbf{q})$ is identically zero. However this cannot happen since Ψ_+ and Ψ_- are eigenfunctions with locally bounded potentials $V(\mathbf{q})$ and therefore satisfy a unique continuation theorem[17]. That is to say if Ψ_+ vanishes in some open set, it vanishes in all of $\mathbf{R}^\nu$ contrary to our original assumption that Ψ_+ is not identically zero. **Q.E.D.**

Remarks.—(1) Both Theorems 1 and 2 continue to be valid if any *attractive* Hubbard model terms are added to the Hamiltonian in (1.1). In other words, the electron-phonon model and Hubbard model results do not interfere with each other as long as they are attractive.

(2) The SSH model is shown here to have at least one spin-singlet state among its ground states. The ground state is not shown to be unique here. The difficulty for the uniqueness proof enters in the derivative of the matrix W in Eq. (3.18).

IV. DISCUSSION

The results presented in this contribution hold only for an even number of electrons in a finite system. In this case, the only nonanalyticities that can enter in properties of the ground state occur when there is a ground-state level crossing. In the cases where the ground state can be shown to be unique, there are never any level crossings, so that any "transition" of the electron-phonon ground state from a collection of delocalized polarons to a collection of self-trapped polarons is not a sharp transition, but is a smooth crossover. We are not prepared to prove any statements about the thermodynamic limit here.

One can ask if these results will survive if a magnetic field is turned on. The answer in general is no because the variational argument presented in Theorem 1 no longer holds if there are interactions with the local electronic spin [as one would have if one added a Zeeman coupling to the Hamiltonian in Eq. (1.1)]. One can also investigate the effect of a magnetic field on the electronic kinetic energy. The hopping matrix elements are always assumed to be real, and therefore can only incorporate flux phases that correspond to entrapped fluxes that are integral multiples of π. If the hopping matrix elements become complex, the entire methodology incorporated here fails, and there are no rigorous statements that we can make about this case.

In conclusion, we have presented a proof that the ground state of a general class of electron-phonon Hamiltonians must include a state that is a spin singlet. The phonons can interact with the electrons in two different ways—the phonons interact with the local electronic charge and the phonons modulate the electronic hopping integrals. The phonon coordinate dependence of the both electron-phonon couplings and the hopping matrix elements is arbitrary. The phonons can be optical modes or acoustical modes, and can contain anharmonic couplings. The hopping matrix is always assumed to be real and symmetric. The Su-Schrieffer-Heeger model, the Holstein model, and the Hubbard model all fall into this general class. In the case where the hopping matrix contains no phonon coordinate dependence, the lattice is connected, the electron-phonon couplings are independent, and the inverse phonon masses are all positive, the ground state has also been shown to be unique.

Acknowledgments

We are indebted to Jan Philip Solovej for valuable discussions, especially for his help with the appendix. This work was supported by the U. S. National Science Foundation under grant no. PHY90-19433 A03 [EHL], and by the Office of Naval Research under grant no. N00014-93-1-0495 [JKF].

APPENDIX: ABSOLUTE VALUE DECREASES KINETIC ENERGY

We shall prove here that replacing a matrix-valued function, $\Psi(\mathbf{q})$ by its absolute value, $|\Psi(\mathbf{q})| = \sqrt{\Psi^\dagger(\mathbf{q})\Psi(\mathbf{q})}$, decreases each component of the kinetic energy

$$T_j(\Psi) := \int\limits_{\mathbf{R}^\nu} \mathrm{Tr}[\partial_j \Psi^\dagger(\mathbf{q})\partial_j\Psi(\mathbf{q})]\mathrm{d}\mathbf{q}, \qquad (A.1)$$

when Ψ is Hermitian. (Here $\partial_j = \partial/\partial q_j$ and $\dagger$ denotes adjoint.)

Before going into the technicalities, let us give a heuristic discussion to motivate the truth of our assertion. If we write, in Dirac notation,

$$\Psi(\mathbf{q}) = \sum_{\alpha=1}^{d} w_\alpha(\mathbf{q})|f_\alpha(\mathbf{q})\rangle\langle f_\alpha(\mathbf{q})|, \qquad (A.2)$$

where d is the dimension of the matrix (in our proof later we shall generalize to $d=\infty$), the functions w_α are the real eigenvalues and the $|f_\alpha(\mathbf{q})\rangle$ are the $\mathbf{q}$-dependent orthonormal eigenfunctions of the Hermitian matrix $\Psi(\mathbf{q})$. Supposing everything to be differentiable we can compute

$$\partial_j\Psi(\mathbf{q}) = \sum_{\alpha=1}^{d} \partial_j w_\alpha(\mathbf{q})|f_\alpha(\mathbf{q})\rangle\langle f_\alpha(\mathbf{q})| + w_\alpha(\mathbf{q})|f_\alpha(\mathbf{q})\rangle\langle g_\alpha(\mathbf{q})| + w_\alpha(\mathbf{q})|g_\alpha(\mathbf{q})\rangle\langle f_\alpha(\mathbf{q})|, \quad (A.3)$$

where $|g_\alpha(\mathbf{q})\rangle = \partial_j|f_\alpha(\mathbf{q})\rangle$. Since $\langle f_\alpha(\mathbf{q})|f_\beta(\mathbf{q})\rangle = \delta_{\alpha\beta}$, we have that $\langle f_\alpha(\mathbf{q})|g_\beta(\mathbf{q})\rangle + \langle g_\alpha(\mathbf{q})|f_\beta(\mathbf{q})\rangle = 0$. Thus, if we square (A.3) and take the trace we have

$$\mathrm{Tr}[\partial_j\Psi(\mathbf{q})]^2 = \sum_{\alpha=1}^{d}[\partial_j w_\alpha(\mathbf{q})]^2 + 2w_\alpha^2(\mathbf{q})\langle g_\alpha(\mathbf{q})|g_\alpha(q)\rangle - 2\sum_{\alpha=1}^{d}\sum_{\beta=1}^{d} w_\alpha(\mathbf{q})w_\beta(\mathbf{q})|\langle g_\alpha(\mathbf{q})|f_\beta(\mathbf{q})\rangle|^2. \qquad (A.4)$$

¿From this we see that replacing $w_\alpha(\mathbf{q})$ by $|w_\alpha(\mathbf{q})|$ can only decrease the last term on the right. The second term does not change. The first term also does not change since, for any real, differentiable function $w(\mathbf{q})$, it is a fact that $[\partial_j w(\mathbf{q})]^2 = [\partial_j|w(\mathbf{q})|]^2$ in the sense of distributions.

Notice that this heuristic discussion gives a pointwise inequality $\mathrm{Tr}[\partial_j\Psi(\mathbf{q})]^2 \geq \mathrm{Tr}[\partial_j|\Psi(\mathbf{q})|]^2$. In our rigorous discussion we shall content ourselves with an inequality for the integral (A.1) — which is sufficient for our purposes in this paper.

We begin the rigorous discussion with some definitions. Let $\mathcal{H}$ be a fixed, separable Hilbert space (for the purposes of our paper $\mathcal{H}$ is finite dimensional, but there is no need for this restriction here). Let $\mathcal{B}$ denote the Hilbert-Schmidt operators on $\mathcal{H}$ and, for $A \in \mathcal{B}$, let

$$\|A\| := \{\mathrm{Tr} A^\dagger A\}^{1/2}$$

denote its *Hilbert-Schmidt norm.* We also define

$$|A| := \sqrt{A^\dagger A} \quad \textit{and} \quad |A^\dagger| := \sqrt{AA^\dagger}$$

and note that $\||A|\| = \||A^\dagger|\| = \|A\|$.

A map

$$\Psi : \mathbf{R}^\nu \to \mathcal{B}$$

is said to be *measurable* if the function $f_A(\mathbf{q}) := \|\Psi(\mathbf{q}) - A\|$ is (Lebesgue) measurable for every Hilbert-Schmidt A. It is not hard to prove that Ψ is measurable if and only if every matrix element $(v, \Psi(\mathbf{q})w)$ is a measurable function for every fixed v and w in $\mathcal{H}$. Ψ is said to be in $L^2(\mathbf{R}^\nu; \mathcal{B})$ if Ψ is measurable and if $f_0 \in L^2(\mathbf{R}^\nu)$, i.e., $\int f_0(\mathbf{q})^2 < \infty$.

The map Ψ is said to be in $H^1(\mathbf{R}^\nu; \mathcal{B})$ if $\Psi \in L^2(\mathbf{R}^\nu; \mathcal{B})$ and if there are maps $\partial_1\Psi, \partial_2\Psi, \ldots, \partial_\nu\Psi$ for which the following holds

(i) Each $\partial_j\Psi : \mathbf{R}^\nu \to \mathcal{B}$ is a map in $L^2(\mathbf{R}^\nu; \mathcal{B})$

(ii) For each infinitely differentiable map, $\phi : \mathbf{R}^\nu \to \mathcal{B}$, of compact support with derivatives $\partial_j\phi$, we have the relation

$$\int_{\mathbf{R}^\nu} \mathrm{Tr}[\partial_j\phi(\mathbf{q})\Psi(\mathbf{q})]\mathrm{d}\mathbf{q} = -\int_{\mathbf{R}^\nu} \mathrm{Tr}[\phi(\mathbf{q})\partial_j\Psi(\mathbf{q})]\mathrm{d}\mathbf{q} \tag{A.5}$$

for each j. (Note: to say that ϕ is differentiable means that for every $\mathbf{q} \in \mathbf{R}^\nu$, $\lim_{\varepsilon\to 0} \|\varepsilon^{-1}[\phi(\mathbf{q}+\varepsilon\mathbf{e}_j) - \phi(\mathbf{q})] - \partial_j\phi(\mathbf{q})\| = 0$ with $\mathbf{e}_j$ being the unit vector in the j^{th} direction. Since ϕ has compact support this limit is uniform in $\mathbf{q}$.) Clearly $\partial_j\Psi^\dagger = (\partial_j\Psi)^\dagger$.

It is easy to verify that $H^1(\mathbf{R}^\nu; \mathcal{B})$ is a Hilbert space with inner product

$$(\Psi, \Psi') = \int_{\mathbf{R}^\nu} \mathrm{Tr}[\Psi^\dagger(\mathbf{q})\Psi'(\mathbf{q}) + \sum_{j=1}^{\nu} \partial_j\Psi^\dagger(\mathbf{q})\partial_j\Psi'(\mathbf{q})]\mathrm{d}\mathbf{q}. \tag{A.6}$$

Clearly, Ψ is in H^1 (or in L^2) if and only if $\Psi^\dagger$ is in H^1 (or in L^2). This class, $H^1(\mathbf{R}^\nu;\mathcal{B})$, is precisely the class needed for quantum mechanics, i.e., so that the variational energy $T_j(\Psi)$ can be defined and so that the norm $\int \mathrm{Tr}\Psi^\dagger\Psi$ can be defined. For matrix–valued functions (i.e., $\mathcal{H}$ is finite dimensional) the properties of measurability, being in L^2 and being in H^1 are just the ordinary meaning of these properties applied to each matrix element of Ψ considered as a function on $\mathbf{R}^\nu$.

We are indebted to Jan Philip Solovej for very considerable help with the following.

THEOREM. *Let Ψ be in $H^1(\mathbf{R}^\nu;\mathcal{B})$. Then $|\Psi|$ and $|\Psi^\dagger|$ are in $H^1(\mathbf{R}^\nu;\mathcal{B})$ and, for each j*

$$2T_j(\Psi) \geq T_j(|\Psi|) + T_j(|\Psi^\dagger|). \tag{A.7}$$

In particular, if $\Psi(\mathbf{q})$ is self-adjoint for all $\mathbf{q}$ then

$$T_j(\Psi) \geq T_j(|\Psi|). \tag{A.8}$$

Remark: If Ψ is not self-adjoint it is quite possible that $T_j(\Psi) < T_j(|\Psi|)$. But then $T_j(\Psi) > T_j(|\Psi^\dagger|)$.

Proof: It suffices to prove the theorem when $\Psi(\mathbf{q})$ is self-adjoint for all $\mathbf{q}$. To see this, consider the Hilbert space $\mathcal{H}_2 = \mathcal{H}\oplus\mathcal{H}$ and replace Ψ by the self adjoint operator $\Psi_2 = \begin{pmatrix} 0 & \Psi^\dagger \\ \Psi & 0 \end{pmatrix}$. Then $|\Psi_2| = \begin{pmatrix} |\Psi| & 0 \\ 0 & |\Psi^\dagger| \end{pmatrix}$ and (A.7) becomes $T_j(\Psi_2) \geq T_j(|\Psi_2|)$.

The first task is to show that $|\Psi|$ is measurable. Our definition of measurability given above is the standard one, namely the inverse image of an open ball in the Banach space $\mathcal{B}$ is measurable. Now the map $\Psi \to |\Psi|$ is continuous in the $\mathcal{B}$ norm and hence, by a standard result, $|\Psi|$ is measurable.

Let $v_1, v_2 \ldots,$ be an orthonormal basis for $\mathcal{H}$. Matrix elements (v_α, Av_β) will be denoted by $A_{\alpha\beta}$. It is easy to verify that $\partial_j(\phi_{\alpha\beta}) = (\partial_j\phi)_{\alpha\beta}$ in the classical sense if ϕ is differentiable. If Ψ has distributional derivatives then $\partial_j(\Psi_{\alpha\beta}) = (\partial_j\Psi)_{\alpha\beta}$ in the *ordinary* distributional sense. [Simply take $\phi(\mathbf{q}) = h(\mathbf{q})|v_\beta\rangle\langle v_\alpha|$ in (A.5), with h being an ordinary infinitely differentiable function of compact support.] Another important preliminary remark is that by Fubini's theorem $\sum_{\alpha,\beta}\int_{\mathbf{R}^\nu} K_{\alpha\beta}(\mathbf{q})\mathrm{d}\mathbf{q} = \int_{\mathbf{R}^\nu}\sum_{\alpha,\beta} K_{\alpha\beta}(\mathbf{q})\mathrm{d}\mathbf{q}$ for any nonnegative, measurable functions $K_{\alpha\beta}(\mathbf{q})$.

For any Ψ in $H^1(\mathbf{R}^\nu;\mathcal{B})$ (including ϕ, as a special case) we define $\partial_{j\varepsilon}\Psi(\mathbf{q})=\varepsilon^{-1}[\Psi(\mathbf{q}+\varepsilon\mathbf{e}_j)-\Psi(\mathbf{q})]$. By the fundamental theorem of calculus for distributions

$$[\partial_{j\varepsilon}\Psi(\mathbf{q})]_{\alpha\beta}=\partial_{j\varepsilon}[\Psi(\mathbf{q})_{\alpha\beta}]=\int_0^1\partial_j[\Psi(\mathbf{q}+t\varepsilon\mathbf{e}_j)_{\alpha\beta}]\mathrm{d}t=\int_0^1[\partial_j\Psi(\mathbf{q}+t\varepsilon\mathbf{e}_j)]_{\alpha\beta}\mathrm{d}t \quad (A.9)$$

Thus, by Fubini's theorem and Schwarz's inequality,

$$\begin{aligned}\int_{\mathbf{R}^\nu}\mathrm{Tr}[\partial_{j\varepsilon}\Psi(\mathbf{q})]^2\mathrm{d}\mathbf{q}&=\sum_{\alpha,\beta}\int_{\mathbf{R}^\nu}\left\{\int_0^1[\partial_j\Psi(\mathbf{q}+t\varepsilon\mathbf{e}_j)]_{\alpha\beta}\mathrm{d}t\right\}^2\mathrm{d}\mathbf{q}\\&\le\sum_{\alpha,\beta}\int_{\mathbf{R}^\nu}\int_0^1[\partial_j\Psi(\mathbf{q}+t\varepsilon\mathbf{e}_j)]^2_{\alpha\beta}\mathrm{d}t\mathrm{d}\mathbf{q}=\int_{\mathbf{R}^\nu}\mathrm{Tr}[\partial_j\Psi(\mathbf{q})]^2\mathrm{d}\mathbf{q}.\end{aligned} \quad (A.10)$$

Therefore, $\partial_{j\varepsilon}\Psi$ is uniformly (in ε) bounded in $L^2(\mathbf{R}^\nu;\mathcal{B})$.

Now we are ready to study $|\Psi|$, and we begin with a little lemma. If N and M are self-adjoint Hilbert-Schmidt operators then

$$\mathrm{Tr}(N-M)^2\ge\mathrm{Tr}(|N|-|M|)^2. \quad (A.11)$$

This is equivalent to $\mathrm{Tr}NM\le\mathrm{Tr}|N||M|$. If we write $2N_\pm=|N|\pm N$ and $2M_\pm=|M|\pm M$, we have that $N_\pm$ and $M_\pm$ are positive semidefinite, and our requirement now reads $\mathrm{Tr}[M_+N_-+M_-N_+]\ge 0$. But this is true because $\mathrm{Tr}M_+N_-=\mathrm{Tr}M_+^{1/2}N_-M_+^{1/2}>0$, etc.

From (A.11) we have that $\mathrm{Tr}[\partial_{j\varepsilon}|\Psi|]^2\le\mathrm{Tr}[\partial_{j\varepsilon}\Psi]^2$, from which we deduce that $\partial_{j\varepsilon}|\Psi|$ is uniformly bounded in $L^2(\mathbf{R}^\nu;\mathcal{B})$. More precisely, by (A.10) and (A.11),

$$\int_{\mathbf{R}^\nu}[\partial_{j\varepsilon}|\Psi|(\mathbf{q})]^2\mathrm{d}\mathbf{q}\le\int_{\mathbf{R}^\nu}\mathrm{Tr}[\partial_j\Psi(\mathbf{q})]^2\mathrm{d}\mathbf{q}. \quad (A.12)$$

Since $L^2(\mathbf{R}^\nu;\mathcal{B})$ is a Hilbert space, the boundedness shown above plus the Banach-Alaoglu theorem implies that there is a sequence $\varepsilon_1,\varepsilon_2,\ldots$, tending to zero such that

$$\partial_{j\varepsilon_n}|\Psi|\rightharpoonup\rho_j, \quad (A.13)$$

as $n\to\infty$, where ρ_j is a map in $L^2(\mathbf{R}^\nu;\mathcal{B})$ and where the convergence is in the weak sense. Since, $\partial_{j\varepsilon}\phi(\mathbf{q})$ converges to $\partial_j\phi(\mathbf{q})$ uniformly (in Hilbert-Schmidt norm) we can write (A.5), using (A.13), as

$$\int_{\mathbf{R}^\nu}\mathrm{Tr}[\phi\rho_j]=\lim_{\varepsilon\to 0}\int_{\mathbf{R}^\nu}\mathrm{Tr}[\phi\partial_{j\varepsilon}|\Psi|]=-\lim_{\varepsilon\to 0}\int_{\mathbf{R}^\nu}\mathrm{Tr}[\partial_{j\varepsilon}\phi|\Psi|]=-\int_{\mathbf{R}^\nu}\mathrm{Tr}[\partial_j\phi|\Psi|] \quad (A.14)$$

The first equality in (A.14) is the weak convergence of $\partial_{j\varepsilon}\Psi$; the second is just a trivial change of variables in the q-integration; the third is the uniform convergence of $\partial_{j\varepsilon}\phi$. Equation (A.14) holds for every ϕ. By uniqueness of the distributional derivative for ordinary functions [and choosing $\phi(\mathbf{q}) = h(\mathbf{q})|v_\beta\rangle\langle v_\alpha|$ as before] we conclude that $(\rho_j)_{\alpha\beta} = (\partial_j|\Psi|)_{\alpha\beta}$, and hence $\rho_j = \partial_j|\Psi|$. However, norms cannot increase under weak limits and thus (A.8) follows from (A.12). **Q.E.D.**

References

[1] A. B. Migdal, Zh. Eksp. Teor. Fiz. **34,** 1438 (1958) [Sov. Phys.—JETP **7,** 999 (1958)].

[2] G. M. Eliashberg, Zh. Eksp. Teor. Fiz. **38,** 966 (1960) [Sov. Phys.—JETP **11,** 696 (1960)].

[3] W. L. McMillan and J. M. Rowell, Ch. 11, *Superconductivity* , edited by R. Parks (Marcel Dekker, Inc., New York, 1969).

[4] P. B. Allen and B. Mitrović, Solid State Phys. **37,** 1 (1982).

[5] J. P. Carbotte, Rev. Mod. Phys. **62,** 1027 (1990).

[6] T. Holstein, Ann. Phys. **8,** 325 (1959).

[7] W. P. Su, J. R. Schrieffer, and A. J. Heeger, Phys. Rev. B **22**, 2099 (1980).

[8] E. H. Lieb, Phys. Rev. Lett. **62,** 1201 (1989).

[9] E. H. Lieb and D. C. Mattis, Phys. Rev. **125,** 164 (1962).

[10] H. Löwen, Phys. Rev. B **37,** 8661 (1988).

[11] L. M. Falicov and J. C. Kimball, Phys. Rev. Lett. **22,** 997 (1969).

[12] T. Kennedy and E. H. Lieb, Physica **138A,** 320 (1986); E. H. Lieb, Physica **140A,** 240 (1986).

[13] E.H. Lieb and M. Loss, Duke Math. J. **71**, 337 (1993).

[14] J.L Lebowitz and N. Macris, *Low temperature phases of itinerant fermions interacting with classical phonons: the static Holstein model*, J. Stat. Phys. **76** to appear (1994).

[15] S. Aubry, G. Abramovici, and J.-L. Raimbault, J. Stat. Phys. **67,** 675 (1992).

[16] J. Hubbard, Proc. Royal Soc. London (Ser. A) **276,** 238 (1963).

[17] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. IV: Analysis of Operators*, (Academic Press, New York, 1978) p. 226.