PHYSICAL REVIEW D VOLUME 15, NUMBER 6 15 MARCH 1977

Fermion-field nontopological solitons*

R. Friedberg
Barnard College and Columbia University, New York, New York 10027

T. D. Lee
Columbia University, New York, New York 10027

(Received 8 December 1976)

The interaction between a scalar field and a set of n fermion fields in three space dimensions is investigated by decomposing the total Hamiltonian H into a sum of two terms: $H = H_{\text{qcl}} + H_{\text{corr}}$, where H_{qcl} denotes the quasiclassical part and H_{corr} the quantum correction. General theorems are given for H_{qcl} concerning the existence of soliton solutions, the general properties of such solutions, and the condition under which the lowest energy state of H_{qcl} is a soliton solution, not the usual plane-wave solution. The effects of the quantum-correction term H_{corr} are examined. It is shown that the quasiclassical solution is a good approximation to the quantum solution over a wide range of the coupling constant. The approximation becomes very good when the fermion number N is large. Even for small N (2 or 3) and weak coupling, the quasiclassical solution remains a fairly good approximation. In the strong-coupling region and for arbitrary N, the quasiclassical approximation becomes again very good, at least when the fermions are nonrelativistic. The question whether the relativistic quantum field theory has a strong-coupling limit or not is not resolved.

I. INTRODUCTION

In this and the subsequent paper of the series, we shall extend our studies of three-space-dimensional nontopological solition solutions[1-4] to include also the fermion field. As before, our interest lies primarily in renormalizable relativistic local quantum field theories.

For definiteness, let us consider the interaction between a single scalar Hermitian field ϕ and a set of Dirac fields ψ^k, where $k=1,2,\ldots,n$. The Hamiltonian density $\mathcal{H}$ is given by

$$\mathcal{H}=\tfrac{1}{2}\Pi^2+\tfrac{1}{2}(\nabla\phi)^2+U(\phi)+\sum_{k=1}^{n}\psi^{k\dagger}(-i\vec{\alpha}\cdot\vec{\nabla}+\beta m+g\beta\phi)\psi^k+\text{counterterms}\ , \tag{1.1}$$

where the dagger denotes Hermitian conjugation, Π and ϕ satisfy the usual commutation relation

$$[\Pi(\vec{r},t),\phi(\vec{r}\,',t)]=-i\delta^3(\vec{r}-\vec{r}\,')\ , \tag{1.2}$$

while ψ^k and $\psi^{k\dagger}$ satisfy the usual anticommutation relations

$$\{\psi^i(\vec{r},t),\psi^k(\vec{r}\,',t)\}=0$$

and (1.3)

$$\{\psi^i(\vec{r},t),\psi^{k\dagger}(\vec{r}\,',t)\}=\delta_{ik}\delta^3(\vec{r}-\vec{r}\,')\ ,$$

and $\vec{\alpha}$ and β are the 4×4 Dirac matrices. Because of renormalizability

$$U(\phi)=\frac{1}{2}a\phi^2+\frac{1}{3!}b\phi^3+\frac{1}{4!}c\phi^4\ . \tag{1.4}$$

The constants a, b, c, m, and g all refer to the appropriate renormalizable constants, and the counterterms in (1.1) are for renormalization purposes (defined entirely in the conventional sense, *without* any consideration of the soliton solutions).

Part of the purpose of this series of papers is to investigate whether all observed hadrons can be regarded as soliton solutions in an appropriate field theory. The above system then serves as a prototype of such models, with ψ^k simulating the "quark" field and $k=1,2,\ldots,n$ the "color" index. Clearly, the Hamiltonian (1.1) is totally unrealistic; other more physical models will be given in the subsequent paper.

The vacuum state is, by definition, the lowest-energy eigenstate of the system when the total fermion number $N=0$. Through the transformation $\phi\to\phi+$ constant, we may always assume for the vacuum state

$$\langle\text{vac}\,|\,\phi(x)\,|\,\text{vac}\rangle=0\ . \tag{1.5}$$

Note that $U(\phi)$ does not contain a linear term in ϕ. However, because of (1.5), there is a linear term in the counterterms. Throughout the paper, we assume the constants

$$a>0,\quad c\geq0,\quad\text{and } b^2\leq 3ac\ , \tag{1.6}$$

so that the absolute minimum of $U(\phi)$ is at $\phi=0$. Also, m is taken to be >0.

In order to find whether the low-lying quantum states consist of solitons or not when the total fermion number $N\neq0$, we introduce an operator $\chi(\vec{r},t)$, defined by

$$\chi(\vec{r},t)\equiv\phi(\vec{r},t)-\phi_c(\vec{r})\ , \tag{1.7}$$

where $\phi_c(\vec{r})$ is a time-independent c-number function, that satisfies

$$\phi_c \to 0 \quad \text{as} \quad r \to \infty \,; \tag{1.8}$$

the detailed form of ϕ_c is yet to be determined. Because of (1.2), we have

$$[\Pi(\vec{r}, t), \chi(\vec{r}', t)] = -i\delta^3(\vec{r} - \vec{r}') \,. \tag{1.9}$$

It is convenient to expand the operator $\psi^k(\vec{r}, t)$ in terms of a complete set of orthonormal c-number time-independent spinor functions $u_l(\vec{r})$ and $v_l(\vec{r})$:

$$\psi^k(\vec{r}, t) = \sum_{l=1}^{\infty} [a_l^k(t) u_l(\vec{r}) + b_l^{k\dagger}(t) v_l(\vec{r})] \,, \tag{1.10}$$

where u_l and v_l are determined by

$$[-i\vec{\alpha} \cdot \vec{\nabla} + \beta(m + g\phi_c)] \times \begin{cases} u_l \\ v_l \end{cases} = \epsilon_l \times \begin{cases} +u_l \\ -v_l \end{cases} \tag{1.11}$$

in which the subscript $l = 1, 2, \ldots \infty$ is arranged so that

$$0 < \epsilon_1 \le \epsilon_2 \le \epsilon_3 \le \cdots \,. \tag{1.12}$$

As will be shown in Appendix A, under rather general conditions, $\epsilon_l \ne 0$. In the standard Dirac representation,[5] $\beta \equiv \rho_3$ and $\vec{\alpha} \equiv \rho_1 \vec{\sigma}$, the spinor $v_l(\vec{r})$ is related to the complex conjugate of the spinor $u_l(\vec{r})$ by $v_l = \rho_2 \sigma_2 u_l^*$. It is clear that ϵ_l, u_l, v_l, a_l^k, and b_l^k are all functionals of ϕ_c. By using (1.3), one sees that the operators a_l^k and b_l^k all anticommute, while

$$\{a_l^k, a_m^{i\dagger}\} = \{b_l^k, b_m^{i\dagger}\} = \delta_{ik}\delta_{lm} \,. \tag{1.13}$$

In terms of χ, Π, a_l^k, and b_l^k, the total Hamiltonian H may be written as a sum of two terms, a quasiclassical part H_{qcl} and a quantum correction H_{corr},

$$H \equiv \int \mathcal{H} d^3r = H_{\text{qcl}} + H_{\text{corr}} \,, \tag{1.14}$$

where $\mathcal{H}$ is given by (1.1),

$$H_{\text{qcl}} = \int [\tfrac{1}{2}(\nabla\phi_c)^2 + U(\phi_c)] d^3r + \sum_{k=1}^{n} \sum_{l=1}^{\infty} \epsilon_l (a_l^{k\dagger} a_l^k + b_l^{k\dagger} b_l^k), \tag{1.15}$$

and

$$H_{\text{corr}} = \int \left\{ \tfrac{1}{2}\Pi^2 + \left[-\nabla^2\phi_c + U'(\phi_c) + g\sum_{1}^{n} \psi^{k\dagger}\beta\psi^k \right]\chi + \tfrac{1}{2}(\nabla\chi)^2 + \tfrac{1}{2}U''(\phi_c)\chi^2 + \frac{1}{3!}U'''(\phi_c)\chi^3 + \frac{1}{4!}c\chi^4 \right\} d^3r - \sum_{1}^{\infty} n\epsilon_l + \text{counterterms}. \tag{1.16}$$

Here, the counterterms refer to exactly the same ones introduced in (1.1), and $U'(\phi) = dU/d\phi$, $U''(\phi) = d^2U/d\phi^2$, etc. The total fermion number operator N is given by

$$N = \sum_{k=1}^{n} \sum_{l=1}^{\infty} (a_l^{k\dagger} a_l^k - b_l^{k\dagger} b_l^k). \tag{1.17}$$

In the quasiclassical Hamiltonian H_{qcl}, ϕ_c is a c-number function, but a_l^k and b_l^k are quantum operators. Thus, the particle-antiparticle symmetry holds for H_{qcl}, as well as for the total Hamiltonian H. Because of the particle-antiparticle conjugation symmetry, only states with $N \ge 0$ will be discussed. Furthermore, keeping in mind the eventual application to hadrons, we shall restrict our subsequent discussion to the sector[6] in which

$$N \le n. \tag{1.18}$$

The lowest eigenvalue E of H_{qcl} can then be derived by distributing the N fermions to the same spinor state $u_1(\vec{r})$ of (1.11), but with different "colors"; consequently, E is the minimum of

$$E(\phi_c) = N\epsilon + \int [\tfrac{1}{2}(\nabla\phi_c)^2 + U(\phi_c)] d^3r, \tag{1.19}$$

where $\epsilon = \epsilon_1(\phi_c)$ as given by (1.11) and (1.12). By labeling the corresponding spinor $u_1 \equiv \psi_c$, we find that the minimum occurs when

$$-\nabla^2\phi_c + U'(\phi_c) = -gN\psi_c^\dagger\beta\psi_c$$

and

$$[-i\vec{\alpha} \cdot \vec{\nabla} + \beta(m + g\phi_c)]\psi_c = \epsilon\psi_c, \tag{1.20}$$

where $\epsilon > 0$ and $\int \psi_c^\dagger \psi_c d^3r = 1$. As before,[1-4] we *define* the soliton solution to be one in which both ϕ_c and ψ_c are permanently confined in space.[7] Of course, (1.20) always has the usual plane-wave solutions, in which $\phi_c = 0$, ψ_c proportional to $\Omega^{-1/2}\exp(i\vec{p} \cdot \vec{r})$, $\epsilon = (m^2 + \vec{p}^2)^{1/2}$, and $\Omega \to \infty$ is the volume of the entire system. As we shall see, besides these plane-wave solutions, the above quasiclassical Hamiltonian H_{qcl} also admits soliton solutions. Because of the boundary condition (1.8), these solitons are of nontopological origin.

In the literature,[8,9] equations identical to (1.20) have been studied for some specific choices of $U(\phi)$ and N; e.g., the well-known SLAC bag model of Bardeen *et al.*[8] is but a special type of soliton solution. However, our concern is of a broader

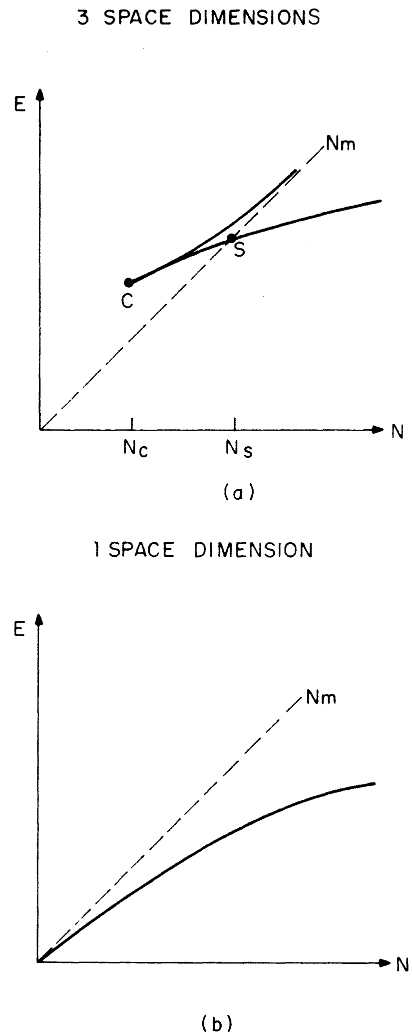

FIG. 1. Schematic drawings of the rest energy E vs the fermion number N, where $N \leq n$ and n = number of fermion fields. The solid curves are for the quasi-classical soliton solutions and the dashed curves for the plane-wave solutions (of zero wave number). Along the solid curve, $\epsilon \equiv dE/dN$ varies continuously from 0 to $m-$; in the three-space-dimensional case, ϵ is also continuous at the spike C.

nature; we are interested in the general conditions under which the low-lying states of $H_{\rm qcl}$ are solitons, *not* the usual plane-wave solutions. The results are stated in a number of theorems given in the next section; these theorems are applicable to an arbitrary $U(\phi)$. Furthermore, in (1.20), the parameter N can be regarded mathematically as a continuous variable. By studying the minimum energy E as a function of N, some new insight into the general character of the soliton solution may then be derived. A schematic drawing of E vs N is given in Fig. 1(a) for the three-space-dimensional case (with details to be given in the next section), and in Fig. 1(b) for the one-space-dimensional case (with details to be given in Appendix C). One sees that for the three-space-dimensional soliton solution, the curve E vs N exhibits a characteristic "spike" shape, similar to the corresponding curve for the boson case.[2,3] There are two critical values N_C and N_S, with $N_S > N_C$. The value N_C gives the location of the spike; for $N < N_C$, there is no soliton solution. The other value N_S is related to the stability point S; for $N > N_S$, the soliton solution has a lower energy than any of the plane-wave solutions. Thus, at least in the quasiclassical approximation, the soliton solution is absolutely stable against decay into plane waves. As will be shown in the next section, it is also stable against fission into several smaller solitons. For $N_S > N > N_C$, the lowest soliton energy is higher than Nm, but nevertheless, it remains stable against infinitesimal perturbations. Furthermore, N_S can be <1, provided that the fermion mass m is sufficiently large; e.g., in the weak-coupling region $(4\pi)^{-1}g^2 \ll 1$ and $c = O(g^2)$, we find $N_S < 1$ if

$$m > \tfrac{1}{2}(\tfrac{3}{2})^6 a^{1/2}(4\pi/g^2). \tag{1.21}$$

For physical applications, N must be an integer. Consequently, if N_S is <1, the lowest-energy state of $H_{\rm qcl}$ changes its character abruptly from the vacuum state ($N=0$) to any other $N \neq 0$ state.

An important question is whether, when one includes the quantum correction $H_{\rm corr}$, the quasi-classical solution remains a good approximation to the exact solution of the total Hamiltonian $H_{\rm qcl} + H_{\rm corr}$. Assuming that the soliton solution is the lowest-energy state of $H_{\rm qcl}$, a relevant parameter is the ratio R between the binding energy (E_b) of the exact solution vs that of the quasiclassical solution:

$$R \equiv (E_b)_{\rm exact}/(E_b)_{\rm qcl}. \tag{1.22}$$

As will be established in Sec. III, and also shown in Table I, in the weak-coupling region $R \to 1$ if $N \gg 1$, and $R \cong 0.768$ if $N = 2$ (assuming, for simplicity, the boson mass $a^{1/2} \ll mg^2/4\pi$). In the strong-coupling region, the exact relativistic quantum solution is not available for comparison. However, we may consider a nonrelativistic fermion model, replacing the Dirac part of the Hamiltonian density, $\psi^{k\dagger}(-i\vec{\alpha}\cdot\nabla + \beta m)\psi^k$ in (1.1), by its nonrelativistic limit $\psi^{k\dagger}(-\frac{1}{2}\nabla^2/m)\psi^k$. In the weak-coupling region, the nonrelativistic model is the limit of the relativistic theory; in the strong-coupling region, the nonrelativistic model gives $R \to 1$ for arbitrary N. Furthermore, when $N \gg 1$, $R \to 1$ for arbitrary coupling. Thus, the quasiclassical solution appears to be a fairly good approximation over a wide range of coupling constants, even when N is small.

The usual perturbation series is an expansion around the plane-wave solutions. When $N > N_S$, the plane-wave solution of $H_{\rm qcl}$ ceases to be the lowest-energy solution; consequently, the convergence of the corresponding perturbation expansion is seriously in doubt. If N_S is <1, this would be the case for all states with $N \neq 0$, which is rather surprising. This somewhat unfamiliar situation is, of course, due to the essential nonlinearity of the problem, and is by no means restricted to field theory, as will be illustrated by the following simple example in elementary mechanics.

Let us consider a single-point particle moving on a plane. The position vector of the particle is $\vec{r} = (x, y)$ and its conjugate momentum is $\vec{p}$. The Hamiltonian is assumed to be

$$\tfrac{1}{2}\vec{p}^{\,2} + \tfrac{1}{2}r^2[(1-gr)^2 + \Delta^2], \tag{1.23}$$

where $r = |\vec{r}|$, and g and Δ are real parameters. The angular momentum $\vec{l} \equiv \vec{r} \times \vec{p}$ is conserved. At a fixed value of $l = |\vec{l}|$, (1.23) becomes

$$\tfrac{1}{2}p_r^{\,2} + V_l(r), \tag{1.24}$$

where p_r is the radial momentum and

$$V_l(r) = \tfrac{1}{2}(l/r)^2 + \tfrac{1}{2}r^2[(1-gr)^2 + \Delta^2]. \tag{1.25}$$

For l and Δ both not too large, $V_l(r)$ has two local minima, say at $r = r_1$ and r_2 with $r_1 < r_2$. As $l \to 0$, one sees that $r_1 \to 0$; hence, r_1 denotes the absolute minimum when l is small. It is easy to show that there exists a critical value l_S. For $l > l_S$, the absolute minimum of $V_l(r)$ changes from r_1 to r_2. Now, in a quantum theory, l takes on only integer values. Thus, if l_S is <1, the character of the $l = 0$ state can be drastically different from all $l \neq 0$ states. This is quite analogous to our field-theoretic problem, in which depending on the parameters, the vacuum state ($N = 0$) may also be significantly different from all $N \neq 0$ states.

Throughout the paper, we adopt the natural units $\hbar$ = velocity of light = 1.

II. QUASICLASSICAL SOLUTIONS

In this section, we consider only the quasiclassical Hamiltonian $H_{\rm qcl}$, which is given by (1.15). At a given fermion number N, the lowest energy of its plane-wave solutions is Nm. Throughout our discussions, we assume (1.18) holds, and therefore the minimum energy E of $H_{\rm qcl}$ is determined by (1.19) and (1.20). As discussed in the previous section, for mathematical convenience N may be regarded as a continuous variable, varying from 0 to ∞.

A. Existence of solitons

Theorem 1. There exists a critical value N_S. For $N > N_S$, the lowest-energy state is a soliton, not the plane-wave solution. Furthermore, as $N \to \infty$,

$$E \le \tfrac{4}{3}\pi\sqrt{2}\,N^{3/4}[U(-m/g)]^{1/4}. \tag{2.1}$$

Theorem 2. N_S is <1 if the fermion mass m is greater than a critical value $m_c = m_c(a, b, c, g)$.

In the weak-coupling region $(4\pi)^{-1}g^2 \ll 1$ and $c = O(g^2)$, an upper bound of m_c is given by

$$m_c < \tfrac{1}{2}(\tfrac{3}{2})^6 a^{1/2}(4\pi/g^2). \tag{2.2}$$

Proof of theorems 1 and 2. From (1.19) and (1.20), it follows that

$$E \le N\left(\int \psi_c^\dagger H_F^{\,2}\psi_c d^3r \Big/ \int \psi_c^\dagger\psi_c d^3r\right)^{1/2} + \int [\tfrac{1}{2}(\nabla\phi_c)^2 + U(\phi_c)]d^3r \tag{2.3}$$

for arbitrary ϕ_c and ψ_c, where

$$H_F = -i\vec{\alpha}\cdot\vec{\nabla} + \beta m + g\beta\phi_c. \tag{2.4}$$

To establish theorem 1, we consider a trial function:

$$\phi_c = \begin{cases} -m/g & \text{for } r \le R \\ -(m/g)\exp[-(r-R)/d] & \text{for } r \ge R \end{cases}$$

and

$$\psi_c = \begin{cases} (2\pi R)^{-1/2}(u/r)\sin\epsilon r & \text{for } r \le R \\ 0 & \text{for } r \ge R, \end{cases} \tag{2.5}$$

where in the standard Dirac representation, $\vec{\alpha} = \rho_1\vec{\sigma}$ and $\beta = \rho_3$,

$$u = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{2.6}$$

R and d are parameters, and

$$\epsilon R = \pi. \tag{2.7}$$

By using (1.4), (2.3), and (2.5), we find for arbitrary R and d,

$$E < \frac{N\pi}{R} + \frac{4\pi}{3}U_0R^3 + \pi(d^{-2} + a)\left(\frac{m}{g}\right)^2(R^2 + Rd + \tfrac{1}{2}d^2)d - \tfrac{2}{9}\pi b\left(\frac{m}{g}\right)^3(R^2 + \tfrac{2}{3}Rd + \tfrac{2}{9}d^2)d + \tfrac{1}{24}\pi c\left(\frac{m}{g}\right)^4(R^2 + \tfrac{1}{2}Rd + \tfrac{1}{8}d^2)d, \tag{2.8}$$

where $U_0 = U(-m/g)$. As N increases, the optimal value of R increases, while that of d remains $O(a^{-1/2})$. For R large, we may therefore keep only the first two terms in the sum on the right-hand side of (2.8); their minimum is $\frac{4}{3}N\pi/R$, and it occurs at

$$R=(\tfrac{1}{4}N/U_0)^{1/4}. \tag{2.9}$$

Consequently, for N large,

$$E\le\tfrac{4}{3}\pi\sqrt{2}\,U_0^{1/4}N^{3/4}+O(N^{1/2}). \tag{2.10}$$

Theorem 1 is then proved.

In deriving (2.9), and therefore also the upper bound (2.10), we assume $U_0>0$, as would be the case if $U(\phi)$ has only one absolute minimum at $\phi=0$. A better upper bound may be derived if there are two absolute minima for $U(\phi)$. As will be shown in Appendix B, if, in addition, the value of $m+g\phi$ changes sign between these two absolute minima, then when $N\to\infty$, instead of (2.10),

$$E\to 3a^{1/2}(2\pi/c)^{1/3}N^{2/3}. \tag{2.11}$$

As we shall also see in Appendix B, the minimum-energy solution in this case is similar to the SLAC bag model[8]; although, in the SLAC bag model, N is $O(1)$.

To prove theorem 2, we first establish the existence of a critical value m_c, so that the lowest soliton energy satisfies

$$E<m \text{ when } N=1 \text{ and } m>m_c. \tag{2.12}$$

We choose, instead of (2.5), the following trial function:

$$\phi_c=Km\exp(-\lambda m r)$$

and

$$\psi_c=(\kappa m)^{3/2}\pi^{-1/2}u\exp(-\kappa m r),$$

where u is given by (2.6) and K, λ, and κ are dimensionless parameters. When $N=1$, by using (2.3) we find for arbitrary values of K, λ, and κ

$$E(1)/m<F\equiv\nu+(\pi K^2/\lambda)\times\left(\frac{1}{2}+\frac{1}{2}\frac{a}{\lambda^2m^2}+\frac{4}{81}\frac{bK}{\lambda^2 m}+\frac{c}{192}\frac{K^2}{\lambda^2}\right), \tag{2.13}$$

where $E(1)$ denotes the value of E at $N=1$, and

$$\nu^2=1+\kappa^2+2gK\left(\frac{2\kappa}{2\kappa+\lambda}\right)^3+g^2K^2\left(\frac{\kappa}{\kappa+\lambda}\right)^3.$$

In order to show $E(1)/m<1$, when m is sufficiently large, we may take the limit $m\to\infty$; in this limit, F approaches

$$F_\infty=\nu+(\pi K^2/\lambda)\left(\frac{1}{2}+\frac{c}{192}\frac{K^2}{\lambda^2}\right).$$

Let $\kappa\equiv\xi|K|^\alpha$ and $\lambda\equiv\eta|K|^\alpha$, where $1>\alpha>\frac{1}{2}$. Keeping ξ, η, and α fixed and $\ne 0$, one sees that as $K\to 0$, $F_\infty\to 1$ and

$$\frac{\partial}{\partial K}F_\infty\to g\left(\frac{2\xi}{2\xi+\eta}\right)^3\ne 0.$$

The minimum of F_∞ must then be less than its value at $K=0$, which is 1. Therefore, (2.12) holds.

At arbitrary N, by using (1.19) and (1.20), one can readily verify that

$$\frac{d}{dN}E(N)=\epsilon(N), \tag{2.14}$$

where $E=E(N)$ and $\epsilon=\epsilon(N)$ are both functions of N. According to (2.7) and (2.9), as $N\to\infty$, $\epsilon(N)\to O(N^{-1/4})\to 0$, while at $N=1$, for $m>m_c$, we have $\epsilon(1)<E(1)<m$ because of (2.12) and (1.19). Thus, when ϵ decreases from $\epsilon(1)<m$ to 0, N increases from 1 to ∞. Since, on account of (2.14), $d(E-Nm)/dN=\epsilon-m$, which is negative for ϵ between $\epsilon(1)$ and 0, we find, for $N>1$, $E(N)-Nm<E(1)-m<0$, provided $m>m_c$. In this case, by definition, N_S is <1, and that leads to the first statement of theorem 2.

Next, we examine the weak-coupling region $(4\pi)^{-1}g^2\ll 1$, but assume $(a^{1/2}/m)=O(g^2)$ and $c=O(g^2)$. In accordance with (1.6), b^2 is $O(ag^2)\ll a$. Since the inequality (2.13) holds for arbitrary κ, λ, and K, we may assume κ and λ to be both $O(g^2)$, and $K=O(g^3)$. The function F in (2.13) becomes

$$F=1+\tfrac{1}{2}\kappa^2+gK\left(\frac{2\kappa}{2\kappa+\lambda}\right)^3+\frac{\pi K^2}{2\lambda}\left(1+\frac{a}{\lambda^2m^2}\right)+O(g^6)\ .$$

Thus, neglecting the $O(g^6)$ term, we find $F=1$ [i.e., $E(1)<m$] when

$$\frac{a}{m^2}=h\equiv-\frac{\lambda^3}{\pi K^2}\left[\kappa^2+2gK\left(\frac{2\kappa}{2\kappa+\lambda}\right)^3+\frac{\pi K^2}{\lambda}\right]. \tag{2.15}$$

It is straightforward to verify that h is maximum when

$$\kappa=\lambda=(\tfrac{2}{3})^6g^2/(2\pi)$$

and

$$K=-\lambda^2(\tfrac{3}{2})^3/g\ ;$$

the corresponding maximum value is $h=\lambda^2$. Consequently, when $m=a^{1/2}/\lambda$, $E(1)$ is <1. The proof of theorem 2 is then completed.

B. Soliton solutions when ϵ is near m

From (1.20) and the boundary condition (1.8), one sees that in order to confine ψ_c in space, ϵ must be $<m$. On the other hand, for $\epsilon>m$ (and when the volume of the system $\Omega\to\infty$), (1.20) has only plane-wave solutions. To find the connection between these two types of solutions, we shall investigate the soliton solution in the limit when $\epsilon \to m-$.

But before taking the limit, let us first convert the Dirac equation into a more convenient form. We multiply the lower equation in (1.20) by $H_F = -i\vec{\alpha}\cdot\vec{\nabla}+\beta(m+g\phi_c)$ on the left. This leads to

$$H_F{}^2\psi_c=\epsilon^2\psi_c\ , \qquad (2.16)$$

where, in the representation $\vec{\alpha}=\rho_1\vec{\sigma}$ and $\beta=\rho_3$,

$$H_F{}^2=-\nabla^2+(m+g\phi_c)^2-g\rho_2\vec{\sigma}\cdot(\vec{\nabla}\phi_c)\ . \qquad (2.17)$$

The solution ψ_c may be decomposed in terms of the eigenfunctions $u_\pm$ of ρ_2:

$$\psi_c=\psi_+u_++\psi_-u_-\ , \qquad (2.18)$$

where

$$\rho_2u_\pm=\pm u_\pm\ ,$$

and ψ_+ and ψ_- are both two-component spinors (in the $\vec{\sigma}$ space) that satisfy

$$[-\nabla^2\mp g\vec{\sigma}\cdot(\vec{\nabla}\phi_c)+(m+g\phi_c)^2]\psi_\pm=\epsilon^2\psi_\pm \qquad (2.19)$$

and

$$\epsilon\psi_\pm=(m+g\phi_c\pm\vec{\sigma}\cdot\vec{\nabla})\psi_\mp\ . \qquad (2.20)$$

One sees that (2.20) implies (2.19), but not vice versa.

When ϵ approaches $m-$, it is convenient to define several dimensionless variables:

$$\xi=(m^2-\epsilon^2)^{1/2}/m,\quad \phi_c=-\tfrac{1}{2}\xi^2xm/g,\quad \tau=\xi mr,$$

and (2.21)

$$N^{1/2}\psi_c=(\tfrac{1}{2}am)^{1/2}(\xi/g)yu+O(\xi^2),$$

where u is the spinor given by (2.6) and $y=y(r)$ is assumed to be radially symmetric. Thus, the last equation in (2.21) implies, in the notation of (2.18),

$$\psi_\pm=\psi_\mp[1+O(\xi)]\ .$$

In the limit $\xi\to0+$, by using (2.21), one finds that, to the lowest order in ξ, the upper equation in (1.20) becomes simply

$$x=y^2\ . \qquad (2.22)$$

Likewise, the lower equation in (1.20) [or its equivalent, (2.19) and (2.20)] can be reduced to

$$\frac{1}{\tau^2}\frac{d}{d\tau}\left(\tau^2\frac{dy}{d\tau}\right)-y+y^3=0\ . \qquad (2.23)$$

Both (2.22) and (2.23) hold for an *arbitrary* $U(\phi)$ of the form (1.4). Solutions of (2.23) have been given explicitly in Ref. 2; by using the virial theorem proved there [Eq. (2.49) of Ref. 2], one has

$$\int y^4d^3\tau=4\int y^2d^3\tau. \qquad (2.24)$$

By using (2.21) and $\int\psi_c^\dagger\psi_c d^3r=1$, we find, when $\epsilon\to0+$,

$$N=\tfrac{1}{2}a(g^2m^2\xi)^{-1}\int y^2d^3\tau\ . \qquad (2.25)$$

Similarly, in the same limit, by using (1.19), (2.21), and (2.24), we obtain the difference between E and Nm to be

$$E-Nm=\tfrac{1}{2}Nm\xi^2>0\ . \qquad (2.26)$$

Now, from the discussions given in the previous section, we know that, when $\epsilon\to0+$, $N\to\infty$ and E is $<Nm$. On the other hand, according to (2.25) and (2.26), when $\epsilon\to m-$, N also $\to\infty$, but E is $>Nm$. Thus, when ϵ varies from 0 to m, the curve denoting the soliton energy $E(N)$ vs N must cross the straight line $E=Nm$ at some point, say S; furthermore, since $dE(N)/dN=\epsilon$ is always positive and $N\to\infty$ at both limits $\epsilon\to0+$ and $m-$, there must be (at least) a spike developed at some point, say C, on the $E(N)$ vs N curve, as shown in Fig. 1(a). Similar curves with "spikes" have also been found for the boson problems[2,3]; such features are, therefore, characteristic of all three-space-dimensional nontopological soliton solutions. (See Appendix C for a discussion of the corresponding one-space-dimensional problem.)

C. Stability

To study stability in the quasiclassical approximation, it is convenient to use the positive-definite form (2.3) for the energy. We define

$$f_c\equiv N^{1/2}\psi_c\ ; \qquad (2.27)$$

the soliton energy $E(N)$ is then, because of (1.19) and (2.3), given by the minimum of the functional $G(f_c,\phi_c)$ at a fixed N, where

$$G(f_c,\phi_c)\equiv N^{1/2}\left(\int f_c^\dagger H_F{}^2f_cd^3r\right)^{1/2}+\int[\tfrac{1}{2}(\nabla\phi_c)^2+U(\phi_c)]d^3r\ , \qquad (2.28)$$

and

$$N=\int f_c^\dagger f_cd^3r\ . \qquad (2.29)$$

For simplicity, we assume that the curve $E(N)$ vs N has only one spike, as shown in Fig. 1(a). By following exactly the proof given in Ref. 2 for its theorem 2, but replacing I, A, and B there by N, ϕ_c, and f_c, respectively, we can establish, for $N>N_C$, the stability of the lowest-energy soliton solution against all infinitesimal perturbations. Since $dE(N)/dN=\epsilon(N)$, and since along the entire branch CS in Fig. 1(a)

$$\frac{d\epsilon}{dN}<0\ ,$$

TABLE I. Ratio R between the exact binding energy and the quasiclassical binding energy. Details are given in Sec. III. For $N=2$ and weak coupling, $R=0.768$ is calculated by assuming, for simplicity, the boson mass $a^{1/2} \ll mg^2/(4\pi)$.

coupling	N	$R \equiv (E_b)_{\rm exact}/(E_b)_{\rm qcl}$ nonrelativistic	relativistic
strong	arbitrary	1	not known
weak	$\gg 1$	1	1
weak	2	0.768	0.768

we find

$$E(N_1+N_2) < E(N_1)+E(N_2) \qquad (2.30)$$

for all positive values of N_1 and N_2. Consequently, the soliton is also stable against fission into several smaller ones. For $N_S > N > N_C$, the soliton energy E is $>Nm$; thus, under finite perturbations, even in the quasiclassical approximation, the soliton can decay into plane-wave solutions.

III. QUANTUM CORRECTIONS

So far, we have considered only the quasiclassical Hamiltonian $H_{\rm qcl}$; the quantum correction will be examined in this section. We shall evaluate the binding energy, including the quantum effect, in various limiting cases, obtaining the results shown in Table I.

For clarity of presentation, we shall assume the renormalized coupling constants

$$b=c=0 \ . \qquad (3.1)$$

Thus, (1.15) and (1.16) become

$$H_{\rm qcl} = \int [\tfrac{1}{2}(\nabla\phi_c)^2 + \tfrac{1}{2}\mu^2\phi_c{}^2]\, d^3r + \sum_{k=1}^{n}\sum_{l=1}^{\infty} \epsilon_l (a_l^{k\dagger}a_l^k + b_l^{k\dagger}b_l^k) \ ,$$

$$H_{\rm corr} = \int [\tfrac{1}{2}\Pi^2 + \tfrac{1}{2}(\nabla\chi)^2 + \tfrac{1}{2}\mu^2\chi^2]\, d^3r + \int \left(-\nabla^2\phi_c + \mu^2\phi_c + g\sum_1^n \psi^{k\dagger}\beta\psi^k\right)\chi\, d^3r - n\sum_1^\infty \epsilon_l + \text{counterterms} \ , \qquad (3.2)$$

where the boson mass $\mu = a^{1/2}$.

In the following, we shall divide our discussions into two parts: (A) assuming the fermions to be nonrelativistic, and (B) the full relativistic Hamiltonian (1.1).

A. Nonrelativistic fermions

For nonrelativistic fermions, we replace in (1.1) the energy density of the Dirac field by its nonrelativistic expression:

$$\sum_{k=1}^{n} \psi^{k\dagger}(-\vec{\alpha}\cdot\vec{\nabla} + \beta m + g\beta\phi)\psi^k \to \sum_{k=1}^{n} \psi^{k\dagger}\left(-\frac{\nabla^2}{2m} + g\phi\right)\psi^k \ . \qquad (3.3)$$

The commutation and anticommutation relations (1.2) and (1.3) remain valid, and as in (1.7) we define $\chi \equiv \phi - \phi_c$. However, instead of (1.10) and (1.11) we have

$$\psi^k(\vec{r}, t) = \sum_{l=1}^{\infty} a_l^k(t) u_l(\vec{r}) \ , \qquad (3.4)$$

where $u_l(\vec{r})$ satisfies the Schrödinger equation

$$[-(2m)^{-1}\nabla^2 + g\phi_c]u_l = \epsilon_l u_l \qquad (3.5)$$

in which the eigenvalue ϵ_l has a degeneracy with respect to the spin orientation of u_l. It is convenient to expand the boson field operator $\chi(\vec{r}, t)$ and its conjugate momentum $\Pi(\vec{r}, t)$ in terms of their Fourier components:

$$\chi(\vec{r}, t) = \sum_{\vec{q}} (2\omega_q\Omega)^{-1/2}[c_{\vec{q}}(t)e^{i\vec{q}\cdot\vec{r}} + c_{\vec{q}}^\dagger(t)e^{-i\vec{q}\cdot\vec{r}}]$$

and

$$\Pi(\vec{r}, t) = -i\sum_{\vec{q}} (\tfrac{1}{2}\omega_q/\Omega)^{1/2}[c_{\vec{q}}(t)e^{i\vec{q}\cdot\vec{r}} - c_{\vec{q}}^\dagger(t)e^{-i\vec{q}\cdot\vec{r}}] \ , \qquad (3.6)$$

where $\Omega \to \infty$ is the volume of the system,

$$\omega_q = (q^2+\mu^2)^{1/2} \text{ and } q = |\vec{q}| \ .$$

Thus, $a_l^k(t)$ and $c_{\vec{q}}(t)$ satisfy, as usual,

$$\{a_l^k(t),\ a_{l'}^{k'}(t)^\dagger\} = \delta_{ll'}\delta_{kk'}$$

and

$$[c_{\vec{q}}(t),\ c_{\vec{q}'}(t)^\dagger] = \delta_{\vec{q}\vec{q}'} \ .$$

The fermion number operator N is given by

$$N = \sum_{k=1}^{n}\sum_{l=1}^{\infty} a_l^{k\dagger}a_l^k \ . \qquad (3.7)$$

To avoid ambiguities with ultraviolet divergence, we impose an upper momentum cutoff Λ in the Fourier expansion of the boson field. If one wishes, one may regard the system as consisting simply of electrons and phonons in a lattice of unit size Λ^{-1}. For $N=n=1$, our problem reduces to the well-known polaron problem,[10] except for the relatively minor differences in ω_q and the interaction form factor. [See the discussions given below after (3.49).] With the momentum cutoff Λ, the field operator ϕ in (3.3) and the c-number function ϕ_c in (3.5) are replaced by, respectively,

$$\phi(\vec{r},t)\to\overline{\phi}(\vec{r},t)\equiv\int f(\vec{r}-\vec{r}\,')\phi(\vec{r}\,',t)\,d^3r' \quad (3.8)$$

and

$$\phi_c(\vec{r})\to\overline{\phi}_c(\vec{r})\equiv\int f(\vec{r}-\vec{r}\,')\phi_c(\vec{r}\,')\,d^3r' ,$$

where the form factor $f(\vec{r}-\vec{r}\,')\to\delta^3(\vec{r}-\vec{r}\,')$ if $\Lambda\to\infty$. As an example, we may take $f(\vec{r})$ to be

$$f(\vec{r})=\begin{cases}(4\pi)^{-1}3\Lambda^3 & \text{if } r\le\Lambda^{-1}\\ 0 & \text{if } r>\Lambda^{-1}\end{cases} . \quad (3.9)$$

The total Hamiltonian H can be decomposed into a sum of two terms:

$$H=H_{\rm qcl}+H_{\rm corr} , \quad (3.10)$$

just as in (1.14). However, instead of (3.2), we have

$$H_{\rm qcl}=\int\left[\tfrac{1}{2}(\nabla\phi_c)^2+\tfrac{1}{2}\mu^2\phi_c^{\,2}\right]d^3r+\sum_{k=1}^{n}\sum_{l=1}^{\infty}\epsilon_l a_l^{k\dagger}a_l^k$$

and $\quad (3.11)$

$$H_{\rm corr}=\sum_{\vec{q}}\omega_q c_{\vec{q}}^\dagger c_{\vec{q}}+\int\left((-\nabla^2\phi_c+\mu^2\phi_c)\chi+g\overline{\chi}\sum_{1}^{n}\psi^{k\dagger}\psi^k\right)d^3r ,$$

where

$$\overline{\chi}(\vec{r},t)=\int f(\vec{r}-\vec{r}\,')\chi(\vec{r}\,',t)\,d^3r' . \quad (3.12)$$

For clarity, we have dropped the counterterm in $H_{\rm corr}$, as we should if we were considering the electron-phonon interaction in a crystal. [See the discussions given below after (3.49).] In any case, this is allowed because the divergence has been removed by the momentum cutoff Λ. Notice that in the nonrelativistic fermion model, $b=c=0$ ensures that there is no diagram leading to the renormalization of b, c, and $\mu=a^{1/2}$.

1. *Quasiclassical solution*

As in (1.19), the quasiclassical energy is now the minimum of

$$E(\phi_c)=N\epsilon+\int\tfrac{1}{2}\left[(\nabla\phi_c)^2+\mu^2\phi_c^{\,2}\right]d^3r , \quad (3.13)$$

where $\epsilon=\epsilon(\phi_c)$ denotes the lowest eigenvalue ϵ_1 in (3.5), and N is assumed to be $\le n$, as before. Just as in (1.20), by labeling the corresponding eigenfunction $u_1\equiv\psi_c$, we find that the minimum occurs when

$$\nabla^2\phi_c-\mu^2\phi_c=gN\sigma$$

and $\quad (3.14)$

$$[-(2m)^{-1}\nabla^2+g\overline{\phi}_c]\psi_c=\epsilon\psi_c ,$$

where $\overline{\phi}_c$ is given by (3.8),

$$\sigma(\vec{r})=\int f(\vec{r}-\vec{r}\,')\psi_c^\dagger(\vec{r}\,')\psi_c(\vec{r}\,')\,d^3r' , \quad (3.15)$$

and ψ_c is normalized according to $\int\psi_c^\dagger\psi_c\,d^3r=1$.

(*i*) $(4\pi)^{-1}g^2\ll\Lambda/(Nm)$. In this case, we may take Λ to be ∞; the form factor $f(\vec{r}-\vec{r}\,')$ becomes simply $\delta^3(\vec{r}-\vec{r}\,')$. To simplify our calculations further, we assume, in addition

$$\mu\ll mg^2/(4\pi) . \quad (3.16)$$

Consequently, we may neglect the terms $\frac{1}{2}\mu^2\phi_c^{\,2}$ and $\mu^2\phi_c$ in (3.13) and (3.14), respectively. It is convenient to introduce the dimensionless scaling variables ρ, $A(\rho)$, and $B(\rho)$:

$$\rho\equiv 2mg^2Nr ,$$
$$A\equiv(2mg^3N^2)^{-1}\phi_c , \quad (3.17)$$
$$B\equiv(2mg^2N)^{-3/2}\psi_c .$$

Hence, $A(\rho)$ and $B(\rho)$ satisfy

$$\nabla_\rho^{\,2}A=B^2 ,$$

and $\quad (3.18)$

$$(-\nabla_\rho^{\,2}+A)B=\hat{\epsilon}B ,$$

where $\int B^2d^3\rho=1$ and

$$\epsilon=2mg^4N^2\hat{\epsilon} . \quad (3.19)$$

The coupled equations (3.18) can be solved numerically. For the nodeless radially symmetric solution,

$$\hat{\epsilon}=0.0814/(4\pi)^2 . \quad (3.20)$$

By using (3.13) and (3.18), one finds

$$\hat{\epsilon}=-\tfrac{3}{4}\int(\nabla_\rho A)^2d^3\rho=-3\int(\nabla_\rho B)^2d^3\rho=\tfrac{3}{4}\int AB^2d^3\rho . \quad (3.21)$$

Hence, for the N-fermion state, the binding energy $E(N)-NE(1)$, in the quasiclassical approximation, is

$$(E_b)_{\rm qcl}=-\tfrac{2}{3}(0.0814)(N^2-1)Nm[g^2/(4\pi)]^2 . \quad (3.22)$$

(*ii*) $(4\pi)^{-1}g^2\gg\Lambda/(Nm)$. When g is very large, the size of the fermion orbit becomes comparable to the lattice size Λ^{-1}. The potential energy $g\overline{\phi}_c$ is $O(g^2N^2\Lambda)$, while the fermion kinetic energy is $O(N\Lambda^2/m)$, which is relatively unimportant. Thus, we may neglect $-(2m)^{-1}\nabla^2$ in the equation for ψ_c; (3.13) becomes, then,

$$E(\phi_c)=gN\int f(\vec{r})\phi_c(\vec{r})d^3r+\tfrac{1}{2}\int[(\nabla\phi_c)^2+\mu^2\phi_c{}^2]d^3r\ ,$$

whose minimum is

$$-(8\pi)^{-1}g^2N^2\int\frac{f(\vec{r})f(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}\exp(-\mu|\vec{r}-\vec{r}\,'|)d^3rd^3r'\ . \tag{3.23}$$

2. *Quantum corrections*

In order to evaluate quantum corrections, it is more convenient to rewrite the total Hamiltonian (3.10) as

$$H=H_0+H_1\ , \tag{3.24}$$

where

$$H_0=H_{\rm qcl}+\sum_{\vec{q}}\omega_q c_{\vec{q}}^\dagger c_{\vec{q}}\ , \tag{3.25}$$

$$H_1=\int\left((-\nabla^2\phi_c+\mu^2\phi_c)\chi+g\bar{\chi}\sum_1^n\psi^{k\dagger}\psi^k\right)d^3r\ , \tag{3.26}$$

$H_{\rm qcl}$ is given by (3.11) and $\bar{\chi}$ by (3.12). In the following, we shall regard H_0 as the zeroth-order Hamiltonian, and H_1 the perturbation.[11]

Let $|0\rangle$ be the state that satisfies

$$a_l^k|0\rangle=c_{\vec{q}}|0\rangle=0 \tag{3.27}$$

for all l, k, and $\vec{q}$. Recalling that (3.5) and (3.14) are related through $\epsilon=\epsilon_1$ and $\psi_c=u_1$, the quasiclassical state for a given fermion number N is simply

$$|N\rangle=\prod_{k=1}^N a_1^{k\dagger}|0\rangle\ , \tag{3.28}$$

where $N\le n$, as before. The state $|N\rangle$ clearly satisfies

$$H_0|N\rangle=H_{\rm qcl}|N\rangle=E_0|N\rangle, \tag{3.29}$$

where E_0 is the quasiclassical energy. For $(4\pi)^{-1}g^2\ll\Lambda/(Nm)$ and $\mu\ll mg^2/(4\pi)$, we have, according to (3.17)–(3.21),

$$E_0=-\tfrac{2}{3}(0.0814)N^3m[g^2/(4\pi)]^2. \tag{3.30}$$

For $(4\pi)^{-1}g^2\gg\Lambda/(Nm)$, E_0 is given by (3.23).

From (3.26), one sees that $H_1|N\rangle$ is a linear superposition of states, all of the form

$$|l,k;\vec{q}\rangle\equiv a_l^{k\dagger}a_1^k c_{\vec{q}}^\dagger|N\rangle. \tag{3.31}$$

The corresponding matrix element $\langle l,k;\vec{q}|H_1|N\rangle$ can be readily derived by setting

$$H_1=\begin{cases}0 & \text{if } l=1\\ 0 & \text{if } k>N\\ g\int\bar{\chi}u_l^\dagger u_1 d^3r & \text{otherwise.}\end{cases} \tag{3.32}$$

The matrix element of H_1 is zero for $l=1$, because of the quasiclassical equation (3.14); it is also zero for $k>N$ because of (3.28). Consequently, the first-order perturbation energy is

$$E_1=\langle N|H_1|N\rangle=0, \tag{3.33}$$

and therefore

$$E_0=\langle N|H|N\rangle. \tag{3.34}$$

In the following, higher-order quantum corrections will be discussed for various limiting cases: (i) superstrong coupling $(4\pi)^{-1}g^2\gg\Lambda/(Nm)$, (ii) strong coupling $(4\pi)^{-1}g^2\gg\ln(\Lambda/m)$, or just $\gg 1$, (iii) $N\gg 1$, but arbitrary coupling, and (iv) $N=2$ and weak coupling. As we shall see, in the first three cases, the quasiclassical solution is a very good approximation to the quantum solution; in case (iv), it remains a fairly good approximation.

(i) Superstrong coupling $(4\pi)^{-1}g^2\gg\Lambda/(N/m)$. A lower bound of the ground-state energy of H can be derived by neglecting in H the fermion-kinetic-energy term $\int\sum\psi^{k\dagger}(-\frac{1}{2}\nabla^2/m)\psi^k d^3r$, since it is a positive-definite operator. We may then take a coordinate representation for the N fermions, and regard their position vectors $\vec{r}_1,\ldots,\vec{r}_N$ as pure parameters. The minimum of H, without the fermion kinetic energy, can be readily obtained, because H depends only quadratically on the boson operators. The minimum occurs when $\vec{r}_1=\vec{r}_2=\cdots=\vec{r}_N$, and this leads to (3.23) being the lower-bound of the quantum ground-state energy; we recall that (3.23) is also the quasiclassical energy E_0 in the superstrong-coupling limit. On the other hand, according to (3.34), E_0 is also an upper bound of the ground-state energy. Thus, in the superstrong-coupling limit, the ratio of the quantum ground-state energy to the quasiclassical energy E_0 approaches 1.

(ii) Strong coupling $(4\pi)^{-1}g^2\gg\ln(\Lambda/m)$, or just $\gg 1$. We may expand the ground-state energy $E_{\rm gd}$ formally in powers of H_1:

$$E_{\rm gd}=E_0+E_2+E_4+\cdots\ . \tag{3.35}$$

The second-order perturbation energy is

$$E_2=\sum_{l,k,\vec{q}}(\epsilon_1-\epsilon_l-\omega_q)^{-1}|\langle l,k;\vec{q}|H_1|N\rangle|^2. \tag{3.36}$$

Because of (3.32), $l\ne 1$ in this sum. Also, since the interaction H_1 is spin-independent, there is no spin-flip matrix element. Let Δ be the minimum of $\epsilon_l+\omega_q-\epsilon_1$ for all $l\ne 1$ states with the same spin orientation as that in the initial state $|N\rangle$. In the strong-coupling region, we may neglect μ. If the coupling is not superstrong, (3.17)–(3.22) are applicable; therefore,

$$\Delta=O(N^2g^4m)\ . \tag{3.37}$$

In the sum (3.36), l extends over the excited bound levels as well as the continuum of (3.5). For the bound levels, the integration over q is finite because of the orbit size R. From (3.17), one sees that

$$R^{-1} = O(Ng^2 m). \tag{3.38}$$

Thus, each excited bound level contributes a term $=O(g^2/(R^2\Delta)) = O(g^2 m)$ to the sum. For the continuum and large q, the recoil energy of the fermion is $\sim\frac{1}{2}q^2/m$. Hence, the q integration would diverge logarithmically if there were no momentum cutoff Λ. We find for Λ large:

$$E_2 = -\tfrac{1}{2}(g/\pi)^2 Nm \ln[\Lambda/(m\Delta)^{1/2}] + O(Ng^2 m). \tag{3.39}$$

If the coupling is strong, but not superstrong, then E_0 is $O(N^3 g^4 m)$. Consequently, $E_2/E_0 \ll 1$, at least in the region $(4\pi)^{-1}N^2 g^2 \gg \ln(\Lambda/m)$.

By pure power counting of momenta, one sees that all higher-order perturbations E_4, E_6, etc. are convergent when $\Lambda \to \infty$. For definiteness, let us first examine E_4. As a typical example, we may consider a single fermion making the following sequence of transitions:

$$l = 1 \to l_1, \omega_1 \to l_2, \omega_1, \omega_2 \to l_3, \omega_1 \to l = 1, \tag{3.40}$$

where l_1, ω_1 denotes that the fermion is in the level l_1, plus an additional boson of frequency ω_1, and l_2, ω_1, ω_2 denotes that the fermion is in the level l_2, plus two bosons of frequencies ω_1 and ω_2, etc. Because of (3.32), $l_1 \neq 1$ and $l_3 \neq 1$, but l_2 is arbitrary. For $l_2 \neq 1$ and ω_1, ω_2 both large, each energy denominator in the perturbation formula gives a factor of the form

$$\approx [\Delta + \omega_q + (2m)^{-1} q^2]^{-1}, \tag{3.41}$$

and there are three such factors in E_4. To obtain the perturbation energy for the N-fermion system, one has to multiply their product by $Ng^4 \int d^3q_1 d^3q_2/(\omega_1\omega_2)$; this leads to a fourth-order energy $= O(Ng^4 m^2/\Delta) = O(m/N)$. A much larger contribution comes when in (3.40), $l_2 = 1$; in that case, only two of the energy denominators are of the form (3.41), the other one is simply $(\omega_1 + \omega_2)^{-1}$. Hence, it gives a fourth-order energy $= O(mg^2)$.

By using (3.26) and (3.32), one may extend the fourth-order energy calculation to include also diagrams in which two different fermions are excited. One such diagram is obtained from (3.40) by setting $l_2 = 1$ and assigning the first pair of transitions to one fermion and the last pair to another. This simply introduces another factor $N-1$, yielding a fourth-order energy

$$E_4 = O(mNg^2). \tag{3.42}$$

A second diagram of the same order of magnitude is obtained by crossing the boson lines, so that (l_3, ω_1) in (3.40) is replaced by (l_3, ω_2). Any other diagram can be obtained from one of these two without changing the topology, simply by altering the "time ordering" of vertices on two different fermion lines. The result of such an alteration is that all three energy denominators become of the form (3.41). The result differs from (3.42) by a factor $O((m/\Delta)^{1/2}) = O(Ng^2)^{-1} \ll 1$. Thus, (3.42) is the dominant contribution to the fourth-order energy.

Likewise, for $E_6, E_8, \ldots$, the dominant contribution always comes from diagrams in which two or more fermions are never simultaneously excited. Any other diagram can be obtained from one of these by a change in the "time ordering" of the transitions, but without altering the topology of the diagram; such a change merely introduces an extra factor $O((m/\Delta)^{1/2}) = O(Ng^2)^{-1} \ll 1$, as in the above fourth-order calculation.

In each of these dominant diagrams, the sequence of transitions can be broken down into clusters. In each cluster, a single fermion makes a sequence of $r+1$ transitions from $l = 1 \to l_1 \to \cdots \to l_r \to l = 1$, where $l_i \neq 1$ $(i = 1, \ldots, r)$. This yields a factor

$$\begin{aligned} O(Ng^{r+1}/\Delta^r) &= O(Ng^{r+1}/(mN^2 g^4)^r) \\ &= O((Ng)^{-(r-1)}(mNg^2)^{-r}). \end{aligned} \tag{3.43}$$

(We are grouping together all diagrams obtainable from one another by shifting entire clusters from one fermion line to another.)

The product of (3.43) over all the clusters is to be multiplied by a number of factors d^3q/ω_q and divided by energy denominators $\sum\omega_q$ for the states between clusters, and integrated. The integrals converge, as seen from (3.41), over a region $q \lesssim O((m\Delta)^{1/2}) = O(mNg^2)$. Therefore, the value of this group of diagrams is of order $(Ng)^{-s}(mNg^2)^t$, where s is the sum of $r-1$ over all clusters, and t must be 1 by dimensional analysis. We now observe that the first line of (3.32) holds good as well when there are additional bosons present in the initial and final states; i.e., on account of (3.14) and (3.26), the matrix element of H_1 is zero if the fermion is making an $l = 1 \to l = 1$ transition. This means that clusters with $r = 0$ are forbidden, and hence $s \geq 0$. The dominant diagrams are therefore those in which all clusters have $r = 1$. Thus, $s = 0$, and

$$E_{2j} = O(mNg^2), \quad \text{all } j \geq 2. \tag{3.44}$$

For coupling $(4\pi)^{-1}g^2 \gg \ln(\Lambda/m)$, but not superstrong, E_0 is $O(N^3 g^4 m)$. By using (3.39) and (3.44) we see that for arbitrary N and for all $j \geq 1$,

$$E_{2j}/E_0 \ll 1 . \tag{3.45}$$

Hence, apart from the convergence question of the series $E_2+E_4+\cdots$, the quasiclassical solution is a very good approximation to the quantum solution.

For coupling $(4\pi)^{-1}g^2 \gg 1$, but not greater than $\ln(\Lambda/m)$, the quasiclassical solution remains the same. The second-order energy E_2 is, however, no longer small compared to E_0. The dominant term in (3.39) is $-\frac{1}{2}(g/\pi)^2 Nm \ln(\Lambda/m)$, which is linear in N and is independent of either the fermion level or the mutual interactions between fermions. Thus, it is useful to separate out this particular level-independent part of self-energy per fermion; we define

$$\begin{aligned} E_0' &= E_0 - \tfrac{1}{2}(g/\pi)^2 Nm \ln(\Lambda/m), \\ E_2' &= E_2 + \tfrac{1}{2}(g/\pi)^2 Nm \ln(\Lambda/m), \end{aligned} \tag{3.46}$$

and

$$E_{2j}' = E_{2j} \quad \text{for} \quad j \geq 2 .$$

The ground-state energy E_{gd} can be rewritten as

$$E_{\text{gd}} = E_0' + E_2' + E_4' + \cdots .$$

Clearly, for all $j \geq 1$,

$$E_{2j}'/E_0' \ll 1 ,$$

provided $(4\pi)^{-1}g^2 \gg 1$. Because this rearrangement (3.46) is independent of level excitations, in all other aspects, e.g., scattering form factors, mobility, etc., the quasiclassical solution remains a good approximation to the quantum solution.

(*iii*) $N \gg 1$ *and arbitrary coupling.* From (3.37), one sees that except in the weak-coupling region, $N \gg 1$ implies $\Delta \gg m$. Therefore, in the series expansion (3.35), the dominant contributions to E_{2j} can be derived by following the same argument given above in (ii), which leads to (3.39) and (3.44). Since E_0 is $O(N^3 g^4 m)$, one obtains $E_{2j}/E_0 \ll 1$ for all $j \geq 1$, provided that N is sufficiently large. The quasiclassical solution is, therefore, a good approximation to the quantum solution.

In the weak-coupling region, as $g^2 \to 0$ for fixed N, the orbital size of the fermion $\to \infty$, and therefore its kinetic energy $\to 0$. Thus, the nonrelativistic fermion case is the limit of the fully relativistic problem. To avoid repetition, we shall defer our discussions to the next section, B, when we study the relativistic case. As we shall see, the quasiclassical solution remains a good approximation to the quantum solution.

(*iv*) $N=2$ *and weak coupling.* In the weak-coupling region $(4\pi)^{-1}g^2 \ll 1$, in order to have a soliton solution for $N=2$, μ/m must be $\ll 1$, in accordance with theorem 2, given in Sec. II. To simplify our discussions, we shall assume $\mu = 0$. The interaction energy between two fermions at distance r apart is Coulomb-like, $-(4\pi)^{-1}g^2/r$. The exact binding energy (E_b) to leading order in g^2 is determined by the ground-state solution of the familiar two-body Schrödinger equation:

$$(E_b)_{\text{exact}} = -\tfrac{1}{4} m \left(\frac{g^2}{4\pi}\right)^2 . \tag{3.47}$$

From (3.22), we see that the corresponding quasiclassical binding energy is

$$(E_b)_{\text{qcl}} = -0.326 m \left(\frac{g^2}{4\pi}\right)^2 . \tag{3.48}$$

Their ratio is

$$R = \frac{(E_b)_{\text{exact}}}{(E_b)_{\text{qcl}}} = 0.768 . \tag{3.49}$$

3. *Polaron*

For the electron-phonon interaction in a polar crystal, we have $n=1$, $\omega_q \cong$ constant and the Fourier transform of the interaction form factor $f(\vec{r})$, defined in (3.8), is proportional to q^{-1}, where q is the phonon momentum; in addition, there is a momentum cutoff Λ due to the lattice size. The soliton solution of such a system is the well-known polaron.[10] Because of the factor q^{-1}, for large Λ the second-order perturbation formula, unlike (3.39), is Λ-independent. Let $E(P)$ be the lowest energy of the state $N=1$ and total momentum P. We may expand $E(P)$ in powers of P:

$$E(P) = E(0) + \tfrac{1}{2}P^2/M + \cdots , \tag{3.50}$$

where $E(0)$ is the same E_{gd}, defined in (3.35). We note that $E(0)$ is related to the work function of the electron, and M to its mobility; both are measurable. This explains why in (3.11) there is no counterterm in H_{corr}.

B. Relativistic case

We now turn to the fully relativistic Hamiltonian. Although the quasiclassical solution can exist in both strong- and weak-coupling regions, the exact relativistic quantum solution is available for comparison only in the weak-coupling region. It is convenient to expand $\chi(\vec{r}, t)$ and its conjugate momentum $\Pi(\vec{r}, t)$ in terms of their Fourier components, as in (3.6). Similarly to (3.24), the relativistic Hamiltonian H can also be written as

$$H = H_0 + H_1 , \tag{3.51}$$

where

$$H_0 = H_{\text{qcl}} + \sum_{\vec{q}} \omega_q c_{\vec{q}}^\dagger c_{\vec{q}} , \tag{3.52}$$

$$H_1 = H_{\text{corr}} - \sum_{\vec{q}} \omega_q c_{\vec{q}}^\dagger c_{\vec{q}} , \tag{3.53}$$

and $H_{\rm qcl}$, $H_{\rm corr}$ are given by (3.2). In the following, we shall regard H_0 as the zeroth-order Hamiltonian, and H_1 the perturbation.[11] The unperturbed state $|N\rangle$ for a given fermion number N remains given by (3.28); i.e.,

$$|N\rangle=\prod_1^N a_1^{k\dagger}|0\rangle \ , \tag{3.54}$$

where the state $|0\rangle$ satisfies

$$a_l^k|0\rangle=b_l^k|0\rangle=c_{\vec{q}}|0\rangle=0$$

for all l, k, and $\vec{q}$. Clearly, $|N\rangle$ satisfies, just as in (3.29),

$$\begin{aligned} H_0|N\rangle&=H_{\rm qcl}|N\rangle \\ &=E_0|N\rangle , \end{aligned} \tag{3.55}$$

where E_0 is the quasiclassical energy.

1. Weak coupling and $N \gg 1$

As we shall see, in order that the weak-coupling expansion holds when $N\gg 1$, we need to assume not only $(4\pi)^{-1}g^2\ll 1$, but also

$$(4\pi)^{-1}g^2\ll N^{-1} \ , \tag{3.56}$$

where $N\leq n$, as before. According to theorem 2 of Sec. II, soliton solutions exist in the weak-coupling limit only if $\mu/m\ll 1$. To simplify our discussions, we assume

$$\mu\ll mg^2/(4\pi) \ , \tag{3.57}$$

so that the main features of the relativistic quasiclassical solutions are approximately described by the $\mu=0$ nonrelativistic limit (3.17)–(3.22).

By using (3.2) and (3.53), we may separate H_1 into two terms:

$$H_1=H_i+H_{ii} \ , \tag{3.58}$$

where

$$H_i=-n\sum_1^\infty \epsilon_l+\text{counterterms} \tag{3.59}$$

and

$$\begin{aligned} H_{ii}=\int \left(-\nabla^2\phi_c+\mu^2\phi_c+g\sum_1^n \psi^{k\dagger}\beta\psi^k\right)\chi d^3r \\ +\text{counterterms} \ . \end{aligned} \tag{3.60}$$

The first term H_i denotes the fermion-loop contribution in the presence of ϕ_c. According to (3.17),

$$\frac{g\phi_c}{m}=O(g^4N^2) \text{ and } \frac{|\nabla\phi_c|}{m\phi_c}=O(g^2N) \ , \tag{3.61}$$

which are both $\ll 1$. We may expand the diagonal matrix element of H_i in powers of $g\phi_c$:

$$E_i\equiv\langle N|H_i|N\rangle=nm^4\int\sum_1^\infty K_j d^3r \ , \tag{3.62}$$

where K_j is $\propto(g\phi_c)^j$. Each K_j may, in turn, be further expanded in terms of the gradient operator ∇, and we shall retain in K_j only the lowest-order ∇ term. Because of the counterterms, $K_1=0$, and ϕ_c^2, $(\nabla\phi_c)^2$, ϕ_c^3, ϕ_c^4 are all absent in the sum $\sum K_j$. Hence, $K_2\sim g^2(\nabla^2\phi_c)^2$, $K_3\sim g^3\phi_c(\nabla\phi_c)^2$, $K_4\sim g^4\phi_c^2(\nabla\phi_c)^2$, and $K_j\sim(g\phi_c)^j$ for $j\geq 5$; among these according to (3.61) the largest ones are K_2 and K_3. In (3.62), the integration $m^2\int d^3r$ brings in another factor $\sim(g^2N)^{-3}$. Consequently, the order of magnitude of E_i is $O(nm(Ng^2)^5)$, which is much smaller than $E_0=O(N^3g^4m)$; their ratio is

$$E_i/E_0=O(N^3g^6)\ll 1 \ . \tag{3.63}$$

In this estimation, N is assumed to be of the same order as n; otherwise, E_i/E_0 is $O(nN^2g^6)$.

We now turn to the second term H_{ii} in (3.59). It is convenient to introduce

$$H_k\equiv g\int \psi^{k\dagger}\beta\psi^k\chi d^3r \ ,$$

and

$$H_\phi\equiv\int(-\nabla^2\phi_c+\mu^2\phi_c)\chi d^3r \ ; \tag{3.64}$$

therefore,

$$H_{ii}=H_\phi+\sum_{k=1}^n H_k+\text{counterterms} \ . \tag{3.65}$$

By using the quasiclassical equation (1.20), we obtain

$$\left\langle N\middle|\left(H_\phi+\sum_{k=1}^n H_k\right)\middle|N\right\rangle=0 \tag{3.66}$$

and

$$\left\langle N,\vec{q}\middle|\left(H_\phi+\sum_{k=1}^n H_k\right)\middle|N\right\rangle=0 \ , \tag{3.67}$$

where

$$|N,\vec{q}\rangle\equiv c_{\vec{q}}^\dagger|N\rangle \ . \tag{3.68}$$

From (3.66), one sees that the first-order perturbation energy of H_{ii} is zero. The second-order perturbation energy can be written as

$$E_{ii}=E_{\psi\psi}+E_{\psi\psi'}+E_{\phi\phi}+E_{\phi\psi} \ , \tag{3.69}$$

in which

$$\begin{aligned} E_{\psi\psi}&=\sum_{k=1}^n\sum_{A\neq N}(E_0-E_A)^{-1}|\langle A|H_k|N\rangle|^2 \\ &\quad+\langle N|\text{counterterms}|N\rangle \ , \\ E_{\psi\psi'}&=\sum_{k'\neq k}\sum_{\vec{q}}(-\omega_q)^{-1}\langle N|H_{k'}|N,\vec{q}\rangle\langle N,\vec{q}|H_k|N\rangle \ , \end{aligned} \tag{3.70}$$

$$E_{\phi\phi}=\sum_{\vec{q}}(-\omega_q)^{-1}|\langle N,\vec{q}|H_\phi|N\rangle|^2 \ ,$$

and

$$E_{\phi\psi}=\sum_{k=1}^{n}\sum_{\vec{q}}(-\omega_q)^{-1}(\langle N|H_k|N,\vec{q}\rangle \times\langle N,\vec{q}|H_\phi|N\rangle+\text{c.c.}) ,$$

where the sum $\sum_A$ in $E_{\psi\psi}$ extends over all eigenstates $|A\rangle$ of H_0, provided $|A\rangle\neq|N\rangle$. The eigenvalue of $|A\rangle$ is E_A.

In (3.69), the first term $E_{\psi\psi}$ is exactly the same as the second-order "Lamb-shift" calculation of the ground level of a fermion in an external attractive potential $=\phi_c$. By following the standard arguments,[12] we find

$$E_{\psi\psi}=O(m(Ng^2)^5\ln(1/Ng^2)) , \tag{3.71}$$

which is much smaller than $E_0=O(N^3g^4m)$ in the weak-coupling region. From (3.54), (3.64), and (3.68), it follows that $\langle N,\vec{q}|H_k|N\rangle$ is independent of k for $k\leq N$, and 0 for $k>N$. Thus, by using (3.67) and (3.70), one sees that

$$E_{\psi\psi'}+\frac{N-1}{N}(E_{\phi\phi}+E_{\phi\psi})=0 . \tag{3.72}$$

For $N\gg1$, we have

$$E_{\psi\psi'}+E_{\phi\phi}+E_{\phi\psi}=O(E_0/N) , \tag{3.73}$$

which can be neglected. Combining (3.71) and (3.73), we see that

$$E_{ii}\ll E_0 , \tag{3.74}$$

which together with (3.63), imply that the effect of H_1 can be neglected in the weak-coupling limit, provided $N\gg1$. The quasiclassical solution is, therefore, a very good approximation to the quantum solution.

We note that by using (3.64), (3.70) and by eliminating the operator χ in terms of its propagator, it can be readily verified that

$$E_{\phi\phi}=-\tfrac{1}{2}\int[(\nabla\phi_c)^2+\mu^2\phi_c^{\,2}]d^3r$$

and

$$E_{\phi\psi}=-\int gN\psi_c^\dagger\beta\psi_c\phi_c\,d^3r .$$

Thus, in effect, $E_{\phi\phi}$ and $E_{\phi\psi}$ just cancel the corresponding terms in the zeroth-order energy E_0; these canceled terms are restored by $E_{\psi\psi'}$, but multiplied by $(N-1)/N$.

2. *Weak coupling and N = 2*

For $N=O(1)$, (3.73) is no longer small; the series expansion in which H_1 is treated as a perturbation is not a valid one. Nevertheless, in the weak-coupling region, the nonrelativistic calculation is the limit of the relativistic problem. From (3.49), we observe that, even for $N=2$, the ratio of the exact binding energy to the quasiclassical value is 0.768, which is a fair approximation.

IV. REMARKS

From our discussions, we see that the quasiclassical solution is a fairly good approximation to the exact quantum solution over a wide range of coupling constants, even when N is small (2 or 3). It is important to note that, when N is small, the quasiclassical-soliton description of an N-body bound state is quite different from the usual description in terms of solutions of the Bethe-Salpeter equation (say, under the ladder approximation). Each of the N bodies is "bloblike" in the former description, while "pointlike" in the latter. In the weak-coupling region, the usual Bethe-Salpeter description is exact; otherwise, the soliton description appears to be a better approximation.

Note added in proof. A simpler and more general proof than the one given in Appendix A that the eigenvalue $\epsilon_l\neq0$ in (1.11) has been given by A. Nishimura [Univ. of Tokyo Report No. 275 (unpublished)] and independently by G. C. Wick (private communication). We wish to thank Dr. Nishimura for communicating his result to us before publication.

ACKNOWLEDGMENT

We wish to thank S. Drell and G. C. Wick for helpful discussion.

APPENDIX A: ABSENCE OF ZERO-ENERGY FERMION LEVEL

In this appendix, we show that under rather general conditions, the eigenvalue ϵ_l of the Dirac equation (1.11) is not zero.

We consider first the one-space-dimensional case. The corresponding Dirac equation is

$$[-i\tau_1 d/dx+\tau_3(m+g\phi_c)]\psi=\epsilon\psi , \tag{A1}$$

where τ_1, τ_2, and τ_3 are the standard (2×2) anticommuting Pauli matrices, and, as before, m is assumed to be >0. If $\epsilon=0$, (A1) becomes

$$\tau_2\frac{d\psi}{dx}=-(m+g\phi_c)\psi .$$

Let u_+ and u_- be the two eigenvectors of τ_2:

$$\tau_2u_\pm=\pm u_\pm .$$

In terms of $u_\pm$, ψ may be written as

$$\psi=\psi_+u_++\psi_-u_- ,$$

where ψ_+ and ψ_- are c-number functions that satisfy $d\psi_\pm/dx=\mp(m+g\phi_c)\psi_\pm$; hence,

$$\psi_\pm=\text{const}\times\exp\left[\mp\int_0^x(m+g\phi_c)dx\right] . \tag{A2}$$

For the nontopological soliton, $\phi_c = 0$ at both limits: $x = +\infty$ and $-\infty$. Consequently, ψ_+ diverges at $x = -\infty$, and ψ_- diverges at $x = +\infty$. As a result, $\epsilon \neq 0$.

On the other hand, as is well known,[8,13,14] ϵ can be 0 for the topological soliton, provided that the two limiting values of $(m + g\phi_c)$ at $x = +\infty$ and $-\infty$ are of opposite sign. For example, if when $x \to +\infty$, $(m + g\phi_c) \to m_+ > 0$, and if when $x \to -\infty$, $(m + g\phi_c) \to m_- < 0$, then $\psi = \psi_+ u_+$ is a good solution of (A1) with $\epsilon = 0$.

Next, we consider the three-space-dimensional case. For simplicity, we assume

$$M \equiv (m + g\phi_c) = M(r) \tag{A3}$$

to be radially symmetric. If in (1.11), the eigenvalue $\epsilon_l = 0$, then the corresponding eigenvector ψ satisfies (in the representation $\vec{\alpha} = \rho_1 \vec{\sigma}$ and $\beta = \rho_3$)

$$\rho_2(\vec{\sigma} \cdot \vec{\nabla})\psi = -M\psi \; . \tag{A4}$$

Let us examine the $S_{1/2}$ solution. (Generalizations to other angular-momentum solutions are straightforward.) We may write

$$\psi = \psi_+ u_+ + \psi_- u_-$$
and
$$\psi_\pm = A_\pm(r) + (\vec{\sigma} \cdot \hat{r}) B_\pm(r) \; , \tag{A5}$$

where $A_\pm(r)$, $B_\pm(r)$ are c-number functions, $u_\pm$ satisfies

$$\rho_2 u_\pm = \pm u_\pm \text{ and } \sigma_3 u_\pm = \lambda u_\pm \; ,$$

with the same λ for both u_+ and u_-. (λ can be either 1 or -1.) By using (A4) and (A5), we find that $A_\pm$ and $B_\pm$ satisfy

$$A_\pm'' + \left(\frac{2}{r} - \frac{M'}{M}\right) A_\pm' - M^2 A_\pm = 0$$
and
$$B_\pm'' + \left(\frac{4}{r} - \frac{M'}{M}\right) B_\pm' - \left(M^2 + \frac{3}{r}\frac{M'}{M}\right) B_\pm = 0 \; , \tag{A6}$$

where a prime denotes d/dr.

For the nontopological soliton, when $r \to \infty$, $\phi_c \to 0$, and therefore $M \to m > 0$. Thus, at infinity, $A_\pm$ and $B_\pm$ have each a regular solution $\sim e^{-mr}$ and an irregular solution $\sim e^{mr}$. Near the origin, let us assume that $M \to \text{const} \times r^s$, where $s \geq 0$; then it follows that, as $r \to 0$, $A_\pm \to \text{const} \times r^\alpha$ and $B_\pm \to \text{const} \times r^\beta$, where

$$\alpha = 0, \text{ or } -1 + s$$
and
$$\beta = s, \text{ or } -3 \; . \tag{A7}$$

One of these independent solutions, $\beta = -3$, is always singular at the origin. Thus, in general, ψ cannot be regular at both limits: $r = 0$ and ∞; consequently, in (1.11) the eigenvalue $\epsilon_l \neq 0$.

In contrast, as shown by Jackiw and Rebbi,[14] in the case of the topological soliton, zero-eigenvalue fermion solutions can be found in three space dimensions, just as in one space dimension.

APPENDIX B: WHEN THE MINIMUM OF $U(\phi)$ IS DEGENERATE

In this case, the parameters a, b, and c in (1.4) satisfy $b^2 = 3ac$. Thus, $U(\phi)$ becomes

$$U(\phi) = \frac{1}{4!} c\phi^2(\phi + 2L)^2, \tag{B1}$$

where $L = (3a/c)^{1/2}$. The minimum $U(\phi) = 0$ now occurs at both $\phi = 0$ and $-2L$. There is, of course, a choice (at least on the quasiclassical level) in defining which minimum should be the vacuum state. For definiteness, we choose $\phi = 0$ to be the vacuum state. Let us define, as in Appendix A,

$$M \equiv m + g\phi \; . \tag{B2}$$

In the following, for convenience of notation, we shall omit the subscript c for the quasiclassical solution, and denote simply

$$\phi = \phi_c \text{ and } \psi = \psi_c \; .$$

We first consider the case that the values of M at these two minima are of opposite signs; i.e.,

$$m > 0, \text{ but } m - 2gL < 0 \; . \tag{B3}$$

For arbitrary values of a, c, m, and g, if N is sufficiently large [though $\leq n$, in accordance with (1.18)], the minimum-energy solution of (1.19) is similar to that of the SLAC bag model.[8] [Though, in the SLAC bag model, N is ~ 1.] We can divide the soliton solution into three regions: the inside region $r \leq R - l$, the shell region $R + l \geq r \geq R - l$, and the outside region $r \geq R + l$, where l is $O(a^{-1/2})$. In the inside region ϕ is $\cong -2L$, while in the outside region ϕ is $\cong 0$. When N is $\gg 1$, R becomes $\gg l$. Therefore, in the shell region, ϕ is approximately determined by the one-space-dimensional solution:

$$\phi = L \tanh[\tfrac{1}{2} a^{1/2}(r - R)] - L + O(R^{-1}) \; . \tag{B4}$$

It will be verified later [through the discussion following (B12)], that if one neglects $O(R^{-1})$, the solution ϕ thus constructed does satisfy (1.20); i.e.,

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2 \frac{d\phi}{dr}\right) - \frac{1}{3!} c\phi(\phi + L)(\phi + 2L) = gN\psi^\dagger \beta \psi \; . \tag{B5}$$

For the moment, we shall assume $\phi(r)$ to be given, and proceed to examine the solution for the fermion field ψ:

$$[-i\vec{\alpha} \cdot \vec{\nabla} + \beta M(r)]\psi = \epsilon\psi \; . \tag{B6}$$

For the $S_{1/2}$ state, we define

$$\psi = \begin{pmatrix} G \\ i(\vec{\sigma}\cdot\hat{r})F \end{pmatrix} \frac{s}{r} , \tag{B7}$$

where $\hat{r} = \vec{r}/r$, and s denotes the eigenstate of σ_3; i.e.,

$$s = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} .$$

In the representation $\vec{\alpha} = \rho_1 \vec{\sigma}$ and $\beta = \rho_3$, the radial functions $F(r)$ and $G(r)$ satisfy

$$\begin{aligned} \left(\frac{d}{dr} + \frac{1}{r}\right) F &= (\epsilon - M) G , \\ \left(\frac{d}{dr} - \frac{1}{r}\right) G &= -(\epsilon + M) F . \end{aligned} \tag{B8}$$

For definiteness, we first consider the region $r > \frac{1}{2}R$. (Here, the factor $\frac{1}{2}$ is arbitrary; it can be any constant <1.) According to (B3) and (B4), $M(r)$ varies from a negative value, near $m - 2gL$ at $r = \frac{1}{2}R$, to a positive value m at $r = \infty$. It is convenient to introduce $A(r)$ and $B(r)$, defined by

$$\begin{aligned} F &= C_0 e^{\theta}(1 + A) , \\ G &= C_0 e^{\theta}(1 + B) , \end{aligned} \tag{B9}$$

where C_0 is a constant, and

$$\theta = -\int_R^r M(r')dr' . \tag{B10}$$

Because of (B8), $A(r)$ and $B(r)$ satisfy

$$\begin{aligned} \left(-\frac{d}{dr} + 2M\right)(A - B) &= \left(\frac{1}{r} - \epsilon\right)(2 + A + B) , \\ -\frac{d}{dr}(A + B) &= \left(\frac{1}{r} + \epsilon\right)(A - B) . \end{aligned} \tag{B11}$$

By comparing (B9) with the corresponding one-space-dimensional solution, one sees that for r large $\gtrsim O(R)$, A and B are both small, $O(R^{-1})$. Therefore, $\psi^\dagger \beta \psi = O(R^{-1})$, and $(B4)$ satisfies (B5).

Because of (B4), $M(r)$ changes rapidly only within the transition region $r = R \pm O(a^{-1/2})$. For $r = O(R)$, but outside the transition region, by using the upper equation of (B11) and regarding d/dr as $O(R^{-1})$, we obtain

$$A - B = (rM)^{-1}(1 - \epsilon r) + O(R^{-2}) .$$

For $R \gg a^{-1/2}$, we may interpolate this solution across the transition region, thus, $1 - \epsilon r = 0$ when $M(r) = 0$, and consequently,

$$\epsilon = R^{-1}[1 + O(a^{-1/2}R^{-1})] .$$

The above argument is essentially the same as that given in Ref. 8 for the SLAC bag model. [Note, however, here the values of $M(r)$ at the two minima of $U(\phi)$ need not be of equal magnitude.]

In the inside region, $r \leq \frac{1}{2}R$, the amplitudes of F and G are exponentially small. In the approximation that M is a constant ($M = m - 2gL < 0$), the r dependence of F and G is given by

$$F = \frac{D}{\epsilon + M}(-\lambda \cosh\lambda r + r^{-1}\sinh\lambda r)$$

and

$$G = D \sinh\lambda r , \tag{B12}$$

where D is a constant, $\lambda = (M^2 - \epsilon^2)^{1/2}$, and $\epsilon \cong R^{-1}$ as before.

Neglecting higher-order terms in $O(R^{-1})$, we obtain the energy of the quasiclassical solution

$$E = N/R + \tfrac{1}{3} 8\pi a^{1/2} L^2 R^2 . \tag{B13}$$

Thus, the mimimum E occurs at $R = \frac{1}{2}[3N/(2\pi a^{1/2}L^2)]^{1/3}$, and its value is

$$E = N^{2/3}(18\pi a^{1/2}L^2)^{1/3} , \tag{B14}$$

which reduces to (2.11), because $L = (3a/c)^{1/2}$.

Next, we consider the case that the minimum of $U(\phi)$ remains degenerate at $\phi = 0$ and $-2L$, but, unlike (B3), at one of the minima $\phi = -2L$,

$$M = m - 2gL = 0 ,$$

whereas at the other minimum $\phi = 0$, $M = m > 0$ as before. In this case, an upper bound of E can be obtained by assuming a trial function in which, as in (B4),

$$\phi = L \tanh[\tfrac{1}{2}a^{1/2}(r - R)] - L + O(R^{-1}) , \tag{B15}$$

but

$$\psi = \begin{cases} (2\pi R)^{-1/2}(u/r)\sin\epsilon r & \text{for } r \leq R \\ 0 & \text{for } r \geq 0 , \end{cases} \tag{B16}$$

as in (2.5). This leads to

$$E \leq \pi N/R + \tfrac{1}{3} 8\pi a^{1/2} L^2 R^2 + O(Ra) .$$

The minimum of this upper bound now occurs at $R = \frac{1}{2}[3N/(2a^{1/2}L^2)]^{1/3}$. We find then, instead of (B14),

$$E \leq \pi N^{2/3}(18 a^{1/2} L^2)^{1/3} . \tag{B17}$$

Lastly, we consider the case that the minimum of $U(\phi)$ remains degenerate at $\phi = 0$ and $-2L$, but M is >0 at both minima. In this case, by following the argument given in Sec. II A, one sees that (2.1) remains valid.

APPENDIX C: SOLUTIONS IN ONE SPACE DIMENSION

In this appendix, we discuss the quasiclassical soliton solutions in one space dimension. As in Appendix B, we shall omit the subscript c, and simply denote $\phi = \phi_c$ and $\psi = \psi_c$.

1. Basic equations

In one space dimension, (1.20) becomes

$$-\frac{d^2}{dx^2}\phi+U'(\phi)=-gN\psi^\dagger\tau_3\psi$$

and

$$\left(-i\tau_1\frac{d}{dx}+\tau_3 M\right)\psi=\epsilon\psi\ , \tag{C1}$$

where, as before, τ_1, τ_2, and τ_3 are the three (2×2) anticommuting Pauli matrices,

$$\epsilon>0 \text{ and } M=m+g\phi\ . \tag{C2}$$

The energy of the system is

$$E=N\epsilon+\int_{-\infty}^{\infty}\left[\frac{1}{2}\left(\frac{d\phi}{dx}\right)^2+U(\phi)\right]dx\ . \tag{C3}$$

In accordance with (2.14),

$$dE/dN=\epsilon\ . \tag{C4}$$

As we shall see, (C1) possesses a simple integration constant; by using this integration constant, one can reduce the above set of coupled equations to a *single* second-order differential equation, which is given by (C14) below.

Let u_+ and u_- be the two eigenvectors of τ_2:

$$\tau_2 u_\pm=\pm u_\pm\ , \tag{C5}$$

with the normalization

$$u_+^\dagger u_+=u_-^\dagger u_-=1\ .$$

It is convenient to introduce two c-number functions $\gamma(x)$ and $\eta(x)$, defined by

$$\psi=\left(\frac{\epsilon}{2N}\right)^{1/2}(\gamma u_+ +\eta u_-)\ . \tag{C6}$$

Thus, γ and η satisfy

$$\left(\frac{d}{dx}+M\right)\gamma=\epsilon\eta$$

and (C7)

$$\left(\frac{d}{dx}-M\right)\eta=-\epsilon\gamma\ .$$

We may eliminate η; (C1) can then be written as

$$-\frac{d^2}{dx^2}\phi+U'(\phi)=-g\gamma\left(\frac{d}{dx}+M\right)\gamma$$

and (C8)

$$\left(-\frac{d^2}{dx^2}+M^2-\frac{dM}{dx}\right)\gamma=\epsilon^2\gamma\ .$$

By using (C2) and (C8), we find

$$\frac{d}{dx}\left[\frac{1}{2}\left(\frac{d\phi}{dx}\right)^2+\frac{1}{2}\left(\frac{d\gamma}{dx}\right)^2+V\right]=0 \tag{C9}$$

where

$$V=\tfrac{1}{2}\epsilon^2\gamma^2-U(\phi)-\tfrac{1}{2}(m+g\phi)^2\gamma^2\ . \tag{C10}$$

Since $\phi=\gamma=0$ at $x=\pm\infty$, (C9) leads to

$$\frac{1}{2}\left(\frac{d\phi}{dx}\right)^2+\frac{1}{2}\left(\frac{d\gamma}{dx}\right)^2+V=0\ . \tag{C11}$$

If one wishes, one may eliminate γ instead of η. This would lead to a discussion completely equivalent to the one above, since (C1) is symmetric under the usual parity operation: $x\to -x$, $\phi\to\phi$, and $\psi\to\tau_3\psi$; therefore, $\gamma\to\eta$ and $\eta\to\gamma$.

2. Mechanical analog

There is a simple mechanical analog for this problem. We may consider the planar motion of a point particle of unit mass, at "time" x and "position" $\vec{r}=(\phi,\gamma,0)$, moving in a static "potential" V and a "magnetic field" $\vec{H}=(0,0,\gamma)$; its equation of motion is

$$\ddot{\vec{r}}=-\vec{\nabla}V+g\dot{\vec{r}}\times\vec{H}\ , \tag{C12}$$

where g is the "charge", $\vec{\nabla}$ is the gradient operator with respect to $\vec{r}$, and the overdot denotes the "time derivative" d/dx. In this mechanical analog, the "energy-conservation" law is (C9). Because of (C11), the "trajectory" that we are interested in has a zero total "energy"; i.e.,

$$\tfrac{1}{2}\dot{\vec{r}}^2+V=0\ . \tag{C13}$$

Furthermore, at both "time limits" $x=+\infty$ and $-\infty$, the trajectory approaches $\vec{r}=0$.

By using (C13), one may eliminate the "time variable" x; the trajectory $\phi=\phi(\gamma)$ is then determined by a single second-order differential equation:

$$-2V\phi''=(1+\phi'^2)\{-V_\phi+V_\gamma\phi' \pm g\gamma[-2V(1+\phi'^2)]^{1/2}\}\ , \tag{C14}$$

where $\phi'=d\phi/d\gamma$, $\phi''=d^2\phi/d\gamma^2$, $V_\phi=\partial V/\partial\phi$, $V_\gamma=\partial V/\partial\gamma$, and the $\pm$ sign is equal to the sign of $d\gamma/dx$.

[It is of interest to note that in the case of the scalar-field soliton problem discussed in Ref. 2, in one space dimension, there exists an identical mechanical analog, in which the point particle moves in the same "potential" V, but without the "magnetic field" $\vec{H}$. Instead of (C12), the equation of motion is simply $\ddot{\vec{r}}=-\vec{\nabla}V$, and therefore, the same energy-conservation law (C13) holds.]

3. Solution when $\epsilon\to m-$

When ϵ approaches $m-$, it is convenient to introduce several dimensionless variables, ξ, τ, X, and y, similar to (2.21):

$$\xi=(m^2-\epsilon^2)^{1/2}/m\,,\quad \phi=-\tfrac{1}{2}\xi^2Xm/g\,,$$
$$\tau=\xi mx\,,\ \text{and}\ \gamma=(\tfrac{1}{2}a/g^2)^{1/2}\xi y\ . \tag{C15}$$

By substituting (C15) into (C14), we may determine the dependence of X on y. In powers of ξ, $X(y)$ is given by

$$X=y^2\pm\xi y^2(1-\tfrac{1}{2}y^2)^{1/2}+O(\xi^2)\ , \tag{C16}$$

where the + sign is for $dy/d\tau>0$, and the − sign for $dy/d\tau<0$. By using (C8) and neglecting $O(\xi)$, we find

$$d^2y/d\tau^2-y+y^3=0\ . \tag{C17}$$

The solution that satisfies $y=0$ at $\tau=\pm\infty$ is

$$y=\sqrt{2}\ \mathrm{sech}\tau\ . \tag{C18}$$

Both (C16) and (C18) are valid for an arbitrary $U(\phi)$, provided $a>0$.

Because of (C6) and $\int\psi^\dagger\psi dx=1$, we have

$$N=\tfrac{1}{2}\epsilon\int(\gamma^2+\eta^2)dx.$$

As $\xi\to0$, $\eta\to\gamma$, and consequently

$$N\to\tfrac{1}{2}(a\xi/g^2)\int y^2d\tau\ .$$

By using (C18), we derive

$$N\to2a\xi/g^2\to0\ \text{as}\ \epsilon\to m-\ . \tag{C19}$$

According to (C4), the derivative of the soliton energy E with respect to N is ϵ, which is $<m$; therefore,

$$E-Nm=\int_0^N(\epsilon-m)dN<0\ . \tag{C20}$$

We find then, in the quasiclassical approximation, for arbitrary N and $U(\phi)$, the lowest one-space-dimensional soliton energy $E(N)$ is *always* less than that of the plane-wave solutions.

From (C16) and (C18), one sees that for arbitrary $U(\phi)$, when $\epsilon\to m-$, the soliton solution corresponds to one in which the fermion density $\rho\equiv N\psi^\dagger\psi$ is maximum at the center of the soliton, say $x=0$. By following an argument similar to the one given in Sec. II A for the proof of theorem 1, one expects as $\epsilon\to0$ (i.e., $N\to\infty$) that the soliton energy $E\le2[\pi U(-m/g)]^{1/2}N^{1/2}$, provided that the absolute minimum of $U(\phi)$ is not degenerate; the maximum of the fermion density ρ in such a solution is also at the center of the soliton. A schematic drawing of $E(N)$ is given in Fig. 1(b).

On the other hand, in one space dimension, if

$$U(\phi)=(4!)^{-1}c\phi^2(\phi+2L)^2\ ,$$

where

$$L=(3a/c)^{1/2}=m/g\ ,$$

so that the condition of the SLAC bag model[8] applies, it can be shown that for N sufficiently large, the lowest-energy soliton solution is of the form of a kink-antikink bound state,[15] whose energy E is $<\frac{4}{3}a^{1/2}L^2$, independent of N. Thus, one expects that, in this case, there is a "cusp," say at $N=N_0$, on the curve of the lowest nontopological soliton energy E vs N: for $N<N_0$, the solution corresponds to one in which the maximum of the fermion density ρ is at the center of the soliton, while for $N>N_0$, the maximum of ρ appears at the two ends of the soliton, similar to the SLAC bag model. This particular "cusp" is purely a one-space-dimensional phenomenon, since it depends on the existence of topological solitons and the corresponding soliton-antisoliton bound states.

---

*Research supported in part by the U.S. Energy Research and Development Administration.

[1]T. D. Lee, Phys. Rep. 23C, 254 (1976).

[2]R. Friedberg, T. D. Lee, and A. Sirlin, Phys. Rev. D 13, 2739 (1976).

[3]R. Friedberg, T. D. Lee, and A. Sirlin, Nucl. Phys. B115, 1 (1976); B115, 32 (1976).

[4]T. D. Lee, paper presented at the Symposium on Frontier Problems in High Energy Physics; Columbia University Report No. CO-2271-76 (unpublished).

[5]P. A. M. Dirac, *The Principles of Quantum Mechanics* (Oxford University Press, London, 1930).

[6]In the sector $N\gg n$, the problem reduces to the description of the abnormal nuclear state considered by T. D. Lee and G. C. Wick [Phys. Rev. D 9, 2291 (1974)] and T. D. Lee and M. Margulies [*ibid.* 11, 1591 (1975)]. See also S. A. Chin and J. D. Walecka, Phys. Lett. 52B, 24 (1974).

[7]Our definition of soliton differs from a more narrow one, used in some mathematical and engineering literature. [See, e.g., the review article by A. C. Scott, F. Y. F. Chu, and D. W. McLaughlin, Proc. IEEE 61, 1443 (1973), and the references cited therein.] In that narrow definition, the term soliton is confined only to some extremely specialized nonlinear equations that have solitary wave solutions whose shape and velocity remain unchanged after any collisions. We propose the term "indestructible solitons" for soliton solutions that satisfy such a narrow definition. For example, solitons of sine-Gordon equations are "indestructible," since they remain intact even after a head-on collision between solitons and antisolitons. For physical applications to particle physics, the requirement of "indestructibility" seems to be an unreasonable one, as different from "stability." Indestructibility would exclude all existing four-dimensional relativistic local field theories; it would also fail to agree with the properties of any particles.

[8]W. A. Bardeen, M. S. Chanowitz, S. D. Drell, M. Weinstein, and T. M. Yan, Phys. Rev. D 11, 1094 (1975).

[9]See also P. Vinciarelli, Lett. Nuovo Cimento 4, 905 (1972); M. Creutz, Phys. Rev. D 10, 1749 (1974); M. Creutz and K. S. Soh, *ibid.* 12, 443 (1975); K. Huang and D. R. Stump, *ibid.* 14, 223 (1976); J. Rafelski. Argonne National Laboratory reports (unpublished); R. Giles, Phys. Rev. D 13, 1670 (1976).

[10]As examples of some earlier references to the polaron problem, see L. D. Landau and S. Pekar, Zh. Eksp. Teor. Fiz. 16, 341 (1946); H. Fröhlich, H. Pelzer, and S. Zieman, Philos. Mag. 41, 221 (1950); T. D. Lee, F. E. Low, and D. Pines, Phys. Rev. 90, 297 (1953); T. D. Lee and D. Pines, *ibid.* 92, 883 (1953); R. P. Feynman, *ibid.* 97, 660 (1955).

[11]Readers may be disturbed by this perturbation expansion, because of the apparent lack of translational invariance of the quasiclassical solution. Such worry turns out to be unnecessary, since the full Hamiltonian is, of course, translationally invariant. This particular question will also be discussed in a separate paper. In the meantime, see L. D. Faddeev and V. E. Korpin, Phys. Lett 63B, 435 (1976), and A. Jevichi, IAS Report No. COO-2220-84 (unpublished) for discussions on related subjects.

[12]See, e.g., J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, 1964), and the references given therein.

[13]C. K. Lee (unpublished).

[14]R. Jackiw and C. Rebbi, Phys. Rev. D 13, 3398 (1976).

[15]R. F. Dashen, B. Hasslacher, and A. Neveu, Phys. Rev. D 10, 4114 (1974); 10, 4130 (1974); 11, 3424 (1975).