Spontaneous dimerization in the $S=\frac{1}{2}$ Heisenberg antiferromagnetic chain with competing interactions

F. D. M. Haldane
Department of Physics, University of Southern California, Los Angeles, California 90007
(Received 5 February 1982)

Spontaneous dimerization is found in the $S=\frac{1}{2}$ isotropic Heisenberg antiferromagnetic chain with competing nearest- and next-nearest-neighbor exchange, $J_2/J_1 \gtrsim \frac{1}{6}$, and results from the same umklapp processes that lead to the Néel state when easy-axis exchange anisotropy is present. Spontaneous and externally induced dimerizations are contrasted.

Recently Haldane[1] noted that the complex critical behavior associated with magnetic fields and anisotropies in the ground state of the exactly soluble $S=\frac{1}{2}$ antiferromagnetic Heisenberg chain was fully explained by hitherto-neglected umklapp processes of a type appropriate to spinless fermions. This was later independently noticed by other authors including Black and Emery[2] and den Nijs,[2] who discussed the importance of this process for understanding a variety of related field theories and two-dimensional critical phenomena. Here, I report that the *same* processes allow the currently topical phenomenon of *spontaneous dimerization* in the isotropic antiferromagnet with *competing exchange* ("frustration") to be explained in terms of the by-now familiar one-dimensional (1D) analog of the Kosterlitz-Thouless critical behavior.[2]

For nearest- and next-nearest-neighbor exchange J_1 and J_2 I find an instability against spontaneous dimerization for $J_2 > J_2^c \simeq \frac{1}{6}J_1$. This throws into perspective recent work[3] on the interesting case $J_2=\frac{1}{2}J_1$, where the ground state is constructed from spontaneously dimerized uncorrelated spin dimers.[4]

I consider the $S=\frac{1}{2}$ spin system

$$\mathcal{H}=\sum_{i=1}^{N} J_1(S_i^x S_{i+1}^x + S_i^y S_{i+1}^y + |\Delta| S_i^z S_{i+1}^z) + J_2 \vec{S}_i \cdot \vec{S}_{i+2} \ , \tag{1}$$

where the anisotropy parameter Δ is introduced to facilitate the discussion. The model is usefully described in terms of fermion variables via the Jordan-Wigner transformation[5] $S_n^z=(c_n^\dagger c_n-\frac{1}{2})$, $S_n^+=(-1)^n[\prod_{j<n}(1-2c_j^\dagger c_j)]c_n^\dagger$; $c_{n+N}^\dagger=-(-1)^\xi c_n^\dagger$, $\xi=S^z+\frac{1}{2}N$, corresponds to periodic spin-boundary conditions. The Hamiltonian becomes

$$\begin{aligned}\mathcal{H}=\sum_i & -\tfrac{1}{2}J_1(c_i^\dagger c_{i+1}+\text{H.c.})+J_1|\Delta|(n_i-\tfrac{1}{2})(n_{i+1}-\tfrac{1}{2}) \\ & +J_2\{(n_i-\tfrac{1}{2})(n_{i+2}-\tfrac{1}{2}) \\ & -[c_i^\dagger(n_{i+1}-\tfrac{1}{2})c_{i+2}+\text{H.c.}]\} \ . \end{aligned} \tag{2}$$

I will be interested in $2k_F$ density-wave states with order parameter $g_p=g_{-p}^*=g(pk_F)$, $p=\pm 1$, $k_F=\frac{1}{2}\pi$, where $g(k)=\langle c_k^\dagger c_{k+\pi}\rangle$. Noting that

$$\sum_n (-1)^n\langle c_n^\dagger c_n\rangle = \sum_k g(k) \ ,$$

$$\sum_n (-1)^n\langle c_n^\dagger c_{n+1}\rangle = i\sum_k \sin(k)g(k) \ ,$$

and in spin variables that these, respectively, correspond to *Néel* order $\langle S_n^z\rangle \sim (-1)^n|g|$, or *dimer* order $\langle S_n^+S_{n+1}^- - S_{n-1}^+S_n^-\rangle \sim (-1)^n|g|$ (i.e., *site*- and *bond*-centered density waves), I identify real g_p with the Néel state, and imaginary g_p with the dimer state.

When $\Delta=J_2=0$, the model corresponds to a free fermion system, and I follow the idea that Luther and Peschel[6] (LP) applied to the model with $J_2=0$—that of linearizing the fermion model about the Fermi level to obtain a continuum field theory. When this is carried out carefully, I obtain the effective Hamiltonian

$$\begin{aligned}\mathcal{H}^{\text{eff}}=J_1\int dx & \left(i\sum_p p\psi_p^\dagger\nabla\psi_p\right)+\gamma_1\left(\sum_{pp'}\rho_p\rho_{p'}\right) \\ & +\gamma_2\left(\sum_p e^{i4pk_Fx}(\psi_p^\dagger\nabla\psi_p^\dagger)(\psi_{-p}\nabla\psi_{-p})\right) \ , \end{aligned} \tag{3}$$

where $\gamma_1=2(|\Delta|+2J_2/J_1)$ and $\gamma_2=(|\Delta|-6J_2/J_1)$; $\psi_p^\dagger(x)$, $p=\pm 1$, are independent Fermi fields, and $\rho_p(x)$ is the correponding density operator. γ_2 is the all-important *spinless-fermion umklapp* term, first identified in Ref. 1, but omitted in the original treatment by LP as the Pauli principle implies $\psi_p^\dagger\psi_p^\dagger\psi_{-p}\psi_{-p}=0$, and this was taken as ruling out umklapp processes. (The term $p=p'$ in γ_1 was also omitted by LP; this term represents renormalizations of the Fermi velocity described as Hartree-Fock terms.[7])

When $J_2/J_1=\frac{1}{6}|\Delta|$, the umklapp term γ_2 vanishes, and $\mathcal{H}^{\text{eff}}$ becomes a simple Luttinger model[8] solved

by a Bogoliubov transformation, and characterized by the *correlation exponent* η: as $|n-n'|\to\infty$, $\langle S_n^+ S_{n'}^-\rangle \sim (-1)^n|n-n'|^{-\eta}$, $\langle c_n^\dagger c_{n'}\rangle \sim (-1)^n|n-n'|^{-(\eta+1/4\eta)}$.[6] $\eta=\frac{1}{2}$ for free fermions, and the solution[8] of the Luttinger model when $\gamma_2=0$ gives $\eta=(\frac{1}{4}+\gamma_1/2\pi)^{1/2}$. Isotropy of the spin-correlation functions dictates that η approaches the value $\eta=1$ in the isotropic limit $|\Delta|=1$.[6,9] When $\gamma_2=0$, the Luttinger model approximation (3) gives $\eta=0.82$ when $|\Delta|=1$ (or $\eta=1$ when $|\Delta|=6J_2/J_1=1.76$), indicating that renormalizations due to nonlinear terms other than γ_2 also give quantitative corrections.[10] This suggests that the special line $J_2(\Delta)$ along which γ_2 vanishes deviates from the value $J_2/J_1=\frac{1}{6}|\Delta|$ at larger values of $|\Delta|$. The gapless, *fluid* character of the Luttinger model suggests the term *"spin fluid"* is an appropriate description of the $\gamma_2=0$ spin chain.

When $\gamma_2\neq 0$, I note following Ref. 1 that the umklapp term can be treated by a scaling theory entirely analogous to that used for the umklapp effects in the spin-$\frac{1}{2}$ Fermi gas.[11] The term γ_2 leads to an instability against a $2k_F$ doubly degenerate density-wave state, with spontaneously broken symmetry.[1,2] I identify $\gamma_2>0$ as leading to the Néel state, and $\gamma_2<0$ as leading to the dimer state. [Note that the canonical transformation $\psi_p^\dagger\to\exp(\frac{1}{4}ip\pi)\psi_p^\dagger$, which changes the sign of γ_2, changes g_p to $(i)^p g_p$.] The scaling equations are of a familiar form,[2,11] and involve γ_2 and the correlation exponent $\eta(\gamma_1)$:

$$\frac{d\tilde{\gamma}_2}{d(\ln D)}=2(\eta^{-1}-1)\tilde{\gamma}_2+O(\tilde{\gamma}_2)^3\ ,$$
$$\frac{d\eta^{-1}}{d(\ln D)}=2\tilde{\gamma}_2^2+O(\tilde{\gamma}_2)^4,\quad \tilde{\gamma}_2=\gamma_2/D\ , \tag{4}$$

where D is an ultraviolet cutoff scale or effective bandwidth. The familiar scaling trajectories of these equations are shown in Fig. 1. *When $|\Delta|=1$, symmetry dictates that the starting point—and subsequent evolution—of the scaling trajectories must be identified the critical scaling trajectories* $\eta^{-1}=1+\tilde{\gamma}_2$. For $J_2<J_2^c\simeq\frac{1}{6}J_1$, the system is described by the stable trajectory scaling to the critical point $\gamma_2=0$, $\eta=1$, and the competing interaction J_2 does not change the character of the simple antiferromagnetic case $J_2=0$. For $J_2>J_2^c$, the system must be identified with the *unstable* critical trajectory, leading away from the point $\gamma_2=0$, $\eta=1$ to the strong-coupling dimer-state fixed point. Note that systems where scaling starts near the line of *unstable* fixed points $\eta^{-1}<1$, $\gamma_2\simeq 0$, can be identified with the sine-Gordon (SG) field theory[12] with coupling parameter $\beta^2=8\pi\eta^{-1}$: the *isotropic* dimer state must thus be identified with the limiting case $\beta^2\to 8\pi$ of the SG theory. The dimer gap, order parameter g_p, and inverse correlation length will all initially grow as $(J_2-J_2^c)^{1/2}\exp[-aJ_1/(J_2-J_2^c)]$ for $J_2>J_2^c$, where a is some numerical constant controlled by the cutoff structure. This transition is very similar to that seen in spin-isotropic systems such as the spin-$\frac{1}{2}$ Fermi gas with backscattering[13] and Kondo models[14] as the coupling changes sign.

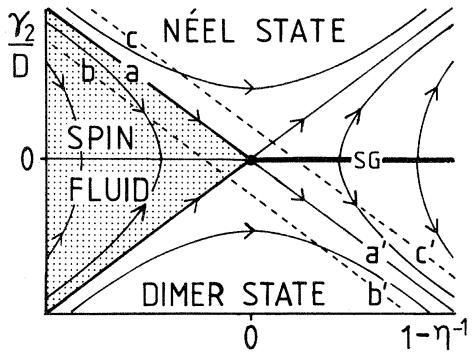

FIG. 1. (See text.) Scaling trajectories of (4): When $|\Delta|=1$, initial parameter values fall on the critical lines aa', scaling either to the limiting critical gapless spin-fluid point $\gamma_2=0$, $\eta=1$ ($J_2<J_2^c$) or to the dimer fixed point ($J_2>J_2^c$). Lines bb' and cc' are the loci of initial values for $|\Delta|\lesssim 1$ and $|\Delta|\gtrsim 1$, respectively. Systems with initial values close to (but not on) the unstable fixed line $\gamma_2=0$, $\eta>1$, are identified with the $\beta^2=8\pi/\eta$ sine-Gordon field theory.

For $|\Delta|<1$, the system will remain in the gapless spin-fluid state that characterizes the planar Heisenberg chain until J_2 exceeds a critical coupling $J_2^c(\Delta)$, when the trajectories will flow to the strong-coupling dimer fixed point. The nature of the transition will now be of "Kosterlitz-Thouless" type,[2] with the order parameter, etc. growing as $\exp\{-b(\Delta)/[J_2-J_2^c(\Delta)]^{1/2}\}$ in the dimer region; the numerical constant $b(\Delta)$ diverges as $|\Delta|\to 1$. For $|\Delta|>1$, $J_2<J_2^c$ there is a similar transition to the Néel state, as seen in the anisotropic chain with $J_2=0$.[1] For $|\Delta|>1$, the two density-wave regions are separated by the gapless line $J_2^c(\Delta)$ along which the umklapp term γ_2 vanishes. Along this line, the Néel and dimer correlations $\langle S_n^z S_{n'}^z\rangle$ and $\langle(\vec{S}_n\cdot\vec{S}_{n+1})(\vec{S}_{n'}\cdot\vec{S}_{n'+1})\rangle$ are the dominant correlations at large separations, both falling off as $(-1)^{(n-n')}|n-n'|^{-(1/\eta)}$, as easily obtained from a Luttinger-model calculation following LP.[6] The critical exponent η^{-1} continuously decreases below 1 along the critical line. Close to this line, the system behaves as a SG system with $\beta^2=8\pi\eta^{-1}$; the principal elementary excitations are solitons carrying $S^z=\pm\frac{1}{2}$ (created in pairs), but in regions adjacent to the section of the critical line with $\eta^{-1}<\frac{1}{2}$, $S^z=0$ "breather" bound-state excitations will also be present ($\beta^2<4\pi$ SG spectrum). The predicted ground-state phase diagram in the $(J_2/J_1, |\Delta|)$ plane is sketched in Fig. 2.

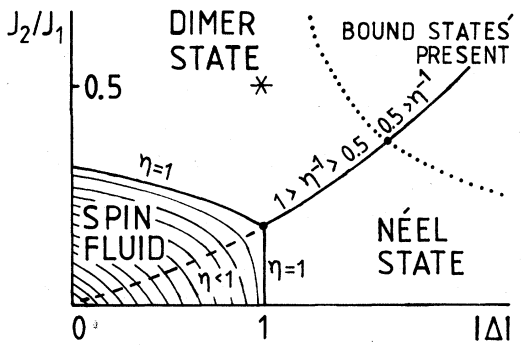

FIG. 2. (Schematic.) Ground-state phase diagram in the $(J_2/J_1, |\Delta|)$ plane. For $|\Delta|<1$, $J_2<J_2^c(\Delta)$, the system is in the gapless *spin-fluid* phase with in-plane spin-correlation exponent $\eta<1$ (lines of constant η are depicted). The umklapp coupling γ_2 vanishes along the broken line, and along its continuation separating the broken-symmetry *dimer* and *Néel* phases, with critical correlation exponent $\eta^{-1}<1$. In these latter phases, the ground state is doublet, with a gap for excitation of *pairs* of $S^z=\pm\frac{1}{2}$ solitons (topological defects); the region where $S^z=0$ breather bound states are present in the gap is shown. The soluble model with $J_2=\frac{1}{2}J_1$ is marked with an asterisk.

In the isotropic model ($|\Delta|=1$), the fundamental excitations in the spontaneously dimerized state are $S=\frac{1}{2}$ soliton states, created only in pairs; the lowest excitations above the doubly degenerate (momenta $P=0, \pm\pi$) ground states of an even-membered ring of spins are thus a continuum of degenerate $S=0$, 1 pair states, with the gap minima at $P=0, \pm\pi$; the identification of the isotropic dimer state with the $\beta^2=8\pi$ SG system rules out breathers or soliton-antisoliton bound states in the gap near $P=0, \pm\pi$. Shastry and Sutherland[3] have recently reported such bound states in a region near the gap *maxima* (at $P=\pm\frac{1}{2}\pi$) for the special model with $J_2/J_1=\frac{1}{2}$; in this range lattice effects are important, and it is outside the scope of the long-wavelength–low-energy SG description used here. The low-energy spectrum deduced here for the dimer state is in complete accord with Ref. 3.

Finally, it is interesting to contrast *spontaneously* dimerized states with those due to an *externally* imposed symmetry-breaking term $\mathcal{H}'=g\sum_n(-1)^n\times\vec{S}_n\cdot\vec{S}_{n+1}$, as considered in the *spin-Peierls* problem.[15] As noted by Cross and Fisher,[15] when translated into fermion variables, this term gives rise to a new SG problem, this time with $\beta^2=2\pi\eta^{-1}$, describing an instability leading to a *singlet* pinned ground state commensurate with the external dimerizing potential, with soliton excitations that now carry $S^z=\pm1$. The isotropic model here corresponds to a $\beta^2=2\pi$ SG system, and the scaling theory shows the dimer gap Δ_d opens as $|g|^{2/3}$.[15] It is interesting to note that $\beta^2=2\pi$ is precisely that value where the SG has just *two* $S^z=0$ breather excitations,[12] with opposite parity, and where the lowest (even parity) breather is precisely degenerate with the $S^z=\pm1$ soliton doublet, forming an $S=1$ triplet; the second (odd parity) breather is a singlet $S=0$ state with a gap $\sqrt{3}\Delta_d$. These two $S=0,1$ states are the only elementary excitations.

If the external dimerizing potential is applied to an *already* spontaneously dimerized isotropic model with $J_2>J_2^c$, a similar spectrum results: The doublet ground-state degeneracy is lifted, and there is now an energy cost linear in the length of regions where the system is in the "wrong" ground state: This imposes a *linear* potential (a 1D Coulomb potential) that confines the $S=\frac{1}{2}$ solitons (i.e., boundaries separating regions of the two now inequivalent dimer configurations) into bound $S=0$ or $S=1$ pairs; the lowest-energy bound state is *symmetric*, with $S=1$.

In the model of the spin-Peierls transition[15] the "external" dimerizing potential arises spontaneously because of lattice distortion; thus topological defects where g changes sign may be "frozen in." For $\beta^2<4\pi$, the energy gain per unit length associated with the opening of the SG gap is finite, and an exact (Bethe ansatz) calculation gives it as $\frac{1}{4}\tan(\frac{1}{2}\pi\kappa)\times\Delta_d^2/v_s$,[16] where v_s is the spin-wave velocity in the limit $g\to0$, and $\kappa=(\beta^2/8\pi)/[1-(\beta^2/8\pi)]=\frac{1}{3}$ when $\beta^2=2\pi$. Since v_s/Δ_d is also the characteristic "healing length" for such a defect (which carries $S=\frac{1}{2}$), the defect energy is of order Δ_d itself. In the absence of interchain coupling, phonon dynamics would allow tunneling motion of the defect, as in recent models of solitons in polyacetylene,[17] and features of the "spontaneously dimerized" spectrum are recovered.

To conclude. The present analysis does *not* explain one interesting feature of the special limit $J_2=\frac{1}{2}J_1$ of the isotropic model—that the correlation between dimers vanishes.[4] However, it places this state in a continuum of spontaneously dimerized states for $J_2>J_2^c\simeq\frac{1}{6}J_1$.

[1]F. D. M. Haldane, Phys. Rev. Lett. 45, 1358 (1980).
[2]J. L. Black and V. J. Emery, Phys. Rev. B 23, 429 (1981); M. P. M. den Nijs, *ibid.* 23, 6111 (1981).
[3]B. S. Shastry and B. Sutherland, Phys. Rev. Lett. 47, 964 (1981).
[4]C. K. Majumdar and D. K. Ghosh, J. Phys. C 3, 911 (1970); J. Math. Phys. 10, 1388, 1399 (1969); P. M. van den Broek, Phys. Lett. 77A, 261 (1980).
[5]P. Jordan and E. Wigner, Z. Phys. 47, 631 (1928).
[6]A. Luther and I. Peschel, Phys. Rev. B 12, 3908 (1975).

[7]H. C. Fogedby, J. Phys. C 11, 4767 (1978).
[8]J. M. Luttinger, J. Math. Phys. 4, 1154 (1963); D. C. Mattis and E. H. Lieb, *ibid.* 6, 304 (1965).
[9]This is directly demonstrated by the derivation of a Luttinger-type representation of form (3) for the isotropic model starting from a *Hubbard chain,* so full spin isotropy is retained throughout the calculation [F. D. M. Haldane, in *Electron Correlation and Magnetism in Narrow-band Systems,* edited by T. Moriya (Springer, Berlin-Heidelberg-New York, 1981), p. 150]. It is also verified by calculation of η from the Bethe ansatz solution of (1) with $J_2=0$, reported in Ref. 1, as well as by LP's original study of the XYZ model gap in Ref. 6.
[10]LP treated the model $J_2=0$ by the linearization scheme, ignoring *both* γ_2 *and* the terms $p=p'$ in γ_1, obtaining the "good" result $\eta\simeq 1.06$ at $|\Delta|=1$. When the γ_1 terms are properly treated, this becomes the not so good $\eta\simeq 0.75$ at $|\Delta|=1$. Omission of the terms γ_2 will always lead to an *under*estimate of η.
[11]V. J. Emery *et al.,* Phys. Rev. B 13, 1262 (1976).
[12]R. F. Dashen *et al.,* Phys. Rev. D 10, 4130 (1974); 11, 3424 (1975).
[13]N. Menyhárd and J. Sólyom, J. Low Temp. Phys. 12, 529 (1973).
[14]P. W. Anderson *et al.,* Phys. Rev. B 1, 4464 (1970).
[15]M. C. Cross and D. S. Fisher, Phys. Rev. B 19, 402 (1979).
[16]F. D. M. Haldane, J. Phys. A 15, 507 (1982).
[17]W. P. Su *et al.,* Phys. Rev. Lett. 42, 1698 (1979); H. Takayama *et al.,* Phys. Rev. B 21, 2388 (1980).