PHYSICAL REVIEW D VOLUME 17, NUMBER 8 15 APRIL 1978

S matrix of the supersymmetric nonlinear σ model

R. Shankar
Department of Physics, Yale University, New Haven, Connecticut

E. Witten
Lyman Laboratory of Physics, Harvard University, Cambridge, Massachusetts 02138
(Received 16 December 1977)

We construct the exact S matrix for the supersymmetric nonlinear σ model in one space and one time dimension. The results confirm that this model possesses mass generation and chiral-symmetry breaking. As a byproduct, we also construct the S matrix for the elementary boson and fermion of the supersymmetric form of the sine-Gordon equation.

I. INTRODUCTION

The two-dimensional nonlinear σ model has attracted considerable attention because of its similarities to four-dimensional gauge theories. The usual σ model, however, is a theory of bosons only, and thus it is really a model not for the full quantum chromodynamics but only for the gauge theory without fermions. It is reasonable to wonder how to include fermions in the σ model so as to obtain a model that would mimic the full quantum chromodynamics with fermions.

In fact, one very natural extension of the nonlinear σ model to include fermions is the supersymmetric model defined by the Lagrangian

$$\mathcal{L}=\frac{1}{g^2}\int d^2x\left[\tfrac{1}{2}(\partial_\mu n^a)^2+\tfrac{1}{2}\bar{\psi}^a i\not{\partial}\psi^a+\tfrac{1}{8}(\bar{\psi}^a\psi^a)^2\right], \qquad (1)$$

and the constraints $n^2=1$, $n\cdot\psi_\alpha=0$; here n^a is a real scalar field of N components, and ψ^a_α is an N-component Majorana Fermi field.[1,2]

[This example suggests a general scheme for the "minimal coupling" of matter fields—fermions or bosons—to the nonlinear σ model. For instance, one could include any number of fields ϕ^i in the vector representation of O(N) with constraints $n^i\phi^i=0$. Or one could include fields ϕ^{ij} in the tensor representation of O(N) with constraints $n^i\phi^{ij}=n^j\phi^{ij}=0$. The general idea is that the matter fields take values tangent to the sphere. The theories defined in this way seem to have many properties in common with four-dimensional gauge theories with matter fields. For instance, the model (1) for $N=3$ has an axial-vector current with an anomalous divergence proportional to the instanton density.[2,3]]

Model (1) is a particularly simple example of fermions interacting with the nonlinear σ model. We would like to answer the following questions about it: Does the model possess dynamical breaking of the discrete γ_5 symmetry and dynamical mass generation, as suggested by the large-N expansion?[4] And is there a special behavior at $N=3$, where the theory possesses instantons and (according to the results of Ref. 1) some additional symmetries?

Recently there has been substantial progress in two-dimensional S-matrix theory—Zamolodchikov,[5] Karowski, Thun, Truong, and Weisz,[6] and Zamolodchikov and Zamolodchikov[7] have shown how to determine the exact S matrices for a number of two-dimensional models, including the sine-Gordon equation, the nonlinear σ model, and the multifermion $(\bar{\psi}\psi)^2$ model. The new methods are applicable to the supersymmetric nonlinear σ model, and we have used them to determine what we believe is the exact S matrix for this model, at least for $N>4$. (As a byproduct we have also determined the S matrix of the supersymmetric form of the sine-Gordon equation.) The results show that the model actually does possess mass generation and symmetry breaking, at least for $N>4$. As we will discuss later, it is not clear that the S matrix we obtain is correct for $N=3$ or 4, so we have not been able to determine what happens in the theory with instantons $N=3$.

In Sec. II we summarize the recent developments in two-dimensional S-matrix theory. In Sec. III we present the S matrix of the supersymmetric nonlinear σ model. Section IV is devoted to a discussion of the properties of the model for small N.

II. THE NEW RESULTS IN TWO-DIMENSIONAL S-MATRIX THEORY

The conserved quantities that we ordinarily encounter in field theory are the momentum P_μ, which transforms as a vector under Lorentz transformations, and internal symmetry charges, which commute with Lorentz transformations.[8]

What distinguishes the sine-Gordon theory and the other theories for which the exact S matrices have been determined is that they possess, in ad-

dition, an infinite number of conserved charges that transform according to higher and higher representations of the Lorentz group—second-rank tensors $Q_{\mu\nu}$, third-rank tensors $Q_{\mu\nu\alpha}$, etc. These charges are all the integrals of local current densities. Moreover, they commute with one another and with the momentum operator P_μ, and therefore also with the mass operator $M^2 = P_\mu P^\mu$. For the sine-Gordon theory there is a conserved tensor of every rank; for the other theories there are an infinite number of conserved tensors of various ranks.

Coleman and Mandula[9] showed that in more than one space dimension a theory possessing conserved charges transforming under the Lorentz group as tensors of second or higher rank would necessarily have a trivial S matrix. In one space dimension, however, this is not true. But the presence of higher conserved charges still places severe restrictions on the S matrix; it is these restrictions that have been used to obtain exact solutions for various two-dimensional models. Here we will summarize these results.

The action of, say, $Q^{\mu\nu\alpha}$ on a one-particle state of momentum p is essentially given by Lorentz invariance as

$$Q^{\mu\nu\alpha}|p\rangle = p^\mu p^\nu p^\alpha |p\rangle . \tag{2}$$

We say "essentially" for two reasons. First there could be terms such as $g^{\mu\nu}p^\alpha$, which are not relevant to the following discussion and will be ignored. Second, since $[Q^{\mu\nu\alpha}, M^2] = 0$, we might have, if there are states degenerate in mass, a structure such as

$$Q^{\mu\nu\alpha}|p^a\rangle = M^{ab} p^\mu p^\nu p^\alpha |p^b\rangle , \tag{3}$$

where $|p^a\rangle$ is a particle of type a with momentum p. We will simply assume $M^{ab} = \delta^{ab}$. Now, the action of $Q^{\mu\nu\alpha}$ on a multiparticle state will be a sum of the action on individual particles since $Q^{\mu\nu\alpha}$ is an integral of a local current and we can always deal with localized, widely separated wave packets. That is,

$$Q^{\mu\nu\alpha}|p_1 p_2 \cdots p_n\rangle = \sum_{i=1}^{n} (p_i^\mu p_i^\nu p_i^\alpha)|p_1, p_2, p_3 \cdots p_n\rangle . \tag{4}$$

Suppose that the state $|p_1 p_2 \cdots p_n\rangle$ scatters to the state $|k_1 k_2 \cdots k_m\rangle$. We know

$$Q^{\mu\nu\alpha}|k_1 k_2 \cdots k_m\rangle = \sum_{j=1}^{m} (k_j^\mu k_j^\nu k_j^\alpha)|k_1 k_2 \cdots k_m\rangle , \tag{5}$$

and, since $Q^{\mu\nu\alpha}$ is conserved, we must have

$$\sum_{i=1}^{n} (p_i^\mu p_i^\nu p_i^\alpha) = \sum_{j=1}^{m} (k_j^\mu k_j^\nu k_j^\alpha) . \tag{6}$$

Considering instead the operator $Q^{\mu\nu\alpha\beta}$, we would get a similar equation with one more Lorentz index:

$$\sum_{i=1}^{n} (p_i^\mu p_i^\nu p_i^\alpha p_i^\beta) = \sum_{j=1}^{m} (k_j^\mu k_j^\nu k_j^\alpha k_j^\beta) . \tag{7}$$

In fact, altogether we get an infinite number of such equations corresponding to the infinite number of conserved tensors of higher rank. This infinite number of equations can be simultaneously satisfied only if $n = m$—so that there is no particle production, no inelasticity—and also if, after suitable relabeling, $p_i = k_i$. Thus the "scattering" consists only of possible time delays and possible exchanges of quantum numbers. N particles entering a collision will emerge later with the same momenta, but possibly after some time delay (or advancement) and possibly with different quantum numbers.

To proceed further, we must understand the action of the $Q^{\mu\nu\alpha}$ on localized wave packets.

For the sake of intuition, we will consider just the purely spatial components Q^{11}, Q^{111}, etc., of our conserved tensors. We let Q^n be the purely spatial component of the nth-rank conserved tensor, and consider the matrix element

$$e^{icQ^n}|p\rangle = e^{icp^n}|p\rangle , \tag{8}$$

where p is the ordinary (spatial) momentum and c is a constant.

We now claim that e^{icQ^n}, acting on a wave packet that is localized in both coordinate and momentum space, will move the packet by an amount dependent on its momentum. In fact, consider a wave packet with wave function

$$\psi(x) = \int_{-\infty}^{\infty} dp\, e^{-a^2(p-p_0)^2} e^{ip(x-x_0)} , \tag{9}$$

where we have taken, for convenience, a Gaussian form of the momentum-space wave function. The operator e^{icQ^n} acting on this state gives a new state with wave function

$$\tilde{\psi}(x) = \int_{-\infty}^{\infty} dp\, e^{-a^2(p-p_0)^2} e^{ip(x-x_0)} e^{icp^n} . \tag{10}$$

To determine where in position space these wave packets are localized, we use stationary phase: The wave packet is concentrated near that value of x where the phase is stationary at $p = p_0$. In this way, we find that $\psi(x)$ is concentrated near $x = x_0$, while $\tilde{\psi}(x)$ is concentrated near $x = x_0 - ncp_0{}^{n-1}$.

Thus, for $n > 1$, the center of the packet is shifted by an amount that depends on its mean momentum p_0. (For $n = 1$, Q' is the ordinary momentum which, of course, moves all packets by the same amount.) Since e^{icQ^n} moves a localized wave packet by an amount dependent on its momentum, it will, when applied to a multiparticle state, with wave packets of different momenta, move them

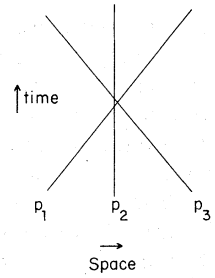

FIG. 1. A three-body collision that cannot be viewed as a sequence of two-body collisions.

relative to one another.

Now, let us consider a collision of three particles of momenta $p_1 < p_2 < p_3$. We will assume that we are dealing with wave packets localized in both position and momentum. A study of the space-time diagrams Figs. 1, 2, and 3 will show that even after the momenta are specified, there are several types of collisions to be considered, depending on the initial positions of the wave packets (or rather, depending on the impact parameters).

In Fig. 1 we show an approximately simultaneous collision of three particles. In Figs. 2 and 3 we show three particles with the same momenta, but with different initial positions, leading to three distinct two-body collisions widely separated in space and time. In Figs. 2 and 3 the chronological sequence of the two collisions is different.

Let us imagine that the two-body scattering amplitudes are known and one wishes to determine the three-body scattering amplitudes. In any theory, a collision of the type in Figs. 2 or 3 can be regarded as a succession of three independent two-body collisions. Its amplitude is the product of the three individual two-body amplitudes.

However, ordinarily we must consider also processess of the type in Fig. 1, which cannot be regarded as the succession of distinct two-body collisions. The amplitude for a collision of the type

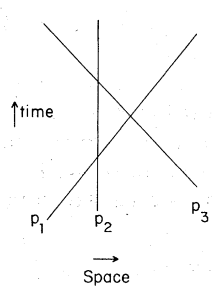

FIG. 2. A three-body collision that can be viewed as a sequence of two-body collisions. The momenta are same as in Fig. 1, but the impact parameters are different.

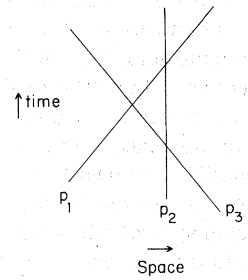

FIG. 3. This collision can also be viewed as a sequence of two-body collisions, but the order of the collisions is different from Fig. 2. In the present theory all three cases are related by symmetry operations and have the same amplitude.

in Fig. 1 is not known even when the two-body amplitudes are known.

But in the theories that we are considering, which possess conserved tensors of various ranks, the situation is much simpler. We simply act on the initial states with one of our symmetry operators e^{icQ^n}. This will, as we have noted, move the three particles relative to one another, thus changing the impact parameters. Since the three diagrams differ only by the different impact parameters, we can, by operating with the e^{icQ^n}, convert any one of the diagrams into any one of the others. Thus the amplitudes for any of the three processes shown are equal.

We learn in this way that any three-body collision can be regarded as a succession of two-body collisions. Schematically, $S^3 = S^2 S^2 S^2$, where S^2 and S^3 are the two- and three-body S matrices. This property is known as the factorization of the three-body S matrix.

Even more, the three-body S matrix can be factored in two ways as a product of two-body S matrices, corresponding to the two possible sequences of collisions indicated in Figs. 2 and 3. Roughly speaking,

$$S^2(p_2,p_3)S^2(p_3,p_1)S^2(p_1,p_2) = S^2(p_1,p_2)S^2(p_1,p_3)S^2(p_2,p_3)\,, \quad (11)$$

where $S^2(p,q)$ is the S matrix for a collision of particles of momentum p and q. This equation is nontrivial because S^2 is a matrix acting on the internal quantum numbers of the particles (which were not indicated in the diagrams) and the different matrices S^2 may not commute. We will refer to (11) as the cubic identity satisfied by the two-particle S matrix. It was a decisive ingredient in the work of Karowski, Thun, Truong, and Weisz,[6] and Zamolodchikov and Zamolodchikov,[7] and we will use it extensively in the next section.

In the next section we will apply the ideas discussed above to determine the S matrix of the supersymmetric nonlinear σ model. In doing so, we assume, of course, that this model possesses conserved tensors of various ranks. This assumption is justified by several considerations. First, the ordinary σ model is known from the solution by Zamolodchikov to possess such conservation laws, and the supersymmetric version of a theory will almost inevitably possess at least all those conversation laws that the original theory possessed. Second, the factorization of the multiparticle S-matrix elements are partly confirmed for this theory by the results of Alvarez for the large-N expansion.

III. THE S MATRIX OF THE SUPERSYMMETRIC σ MODEL

As stated earlier, the model whose S matrix we wish to determine is the supersymmetric nonlinear σ model, described by the Lagrangian

$$\mathcal{L}=\frac{1}{g^2}\int d^2x\left[\tfrac{1}{2}(\partial_\mu n^a)^2+\tfrac{1}{2}\bar{\psi}^a i\not{\partial}\psi^a+\tfrac{1}{8}(\bar{\psi}^a\psi^a)^2\right]. \quad (12)$$

We will assume, in agreement with the results from the large-N expansion,[4] that the spectrum consists of a degenerate supermultiplet of N massive boson states $|b^a\rangle$ and N massive fermion states $|f^a\rangle$, $a=1,2,\ldots,N$; we will work in units in which the mass is equal to one.

This theory contains a conserved Majorana supercharge Q_α. In terms of the chiral components $Q_\pm$ of this supercharge, the supersymmetry algebra is

$$\begin{aligned} &Q_+^{\,2}=P_0+P_1\,,\\ &Q_-^{\,2}=P_0-P_1\,,\\ &Q_+Q_-+Q_-Q_+=0\,,\\ &[Q_\pm,P_\mu]=0\,. \end{aligned} \quad (13)$$

As a kinematical variable, we will, following Zamolodchikov, use the rapidity θ, related to the energy and momentum by $E=\cosh\theta$, $q=\sinh\theta$.

It follows from the algebra (13) that, with suitable choices for the phases of the states, the action of Q_+ and Q_- on one-particle states is

$$\begin{aligned} &Q_+|b^a(\theta)\rangle=e^{\theta/2}|f^a(\theta)\rangle\,,\\ &Q_+|f^a(\theta)\rangle=e^{\theta/2}|b^a(\theta)\rangle\,,\\ &Q_-|b^a(\theta)\rangle=ie^{-\theta/2}|f^a(\theta)\rangle\,,\\ &Q_-|f^a(\theta)\rangle=-ie^{-\theta/2}|b^a(\theta)\rangle\,, \end{aligned} \quad (14)$$

where $|b^a(\theta)\rangle$, for example, is a one-boson state of isospin a and rapidity θ. [In writing (14), we use the fact that Q_+ and Q_- transform under Lorentz transformations like the chiral components of a spinor.]

Now, we would like to write down the most general form for the two-body S matrix that is allowed by supersymmetry, isospin, and Lorentz invariance. Lorentz invariance implies that the matrix is only a function of the rapidity difference between the two particles, so that we can work in the center-of-mass frame and consider a collision of a particle of rapidity $\theta/2$ with one of rapidity $-\theta/2$. The rapidity difference between the two particles is related to the usual kinematical variable $s=(p_1+p_2)^2$ by $s=4\cosh^2(\frac{1}{2}\theta)$.

There are four "channels" to consider—the initial state may be

$$|b^a(\tfrac{1}{2}\theta)b^b(-\tfrac{1}{2}\theta)\rangle,\quad |f^a(\tfrac{1}{2}\theta)f^b(-\tfrac{1}{2}\theta)\rangle,$$
$$|f^a(\tfrac{1}{2}\theta)b^b(-\tfrac{1}{2}\theta)\rangle,\quad \text{or}\quad |b^a(\tfrac{1}{2}\theta)f^b(-\tfrac{1}{2}\theta)\rangle\,.$$

It is easy to diagonalize the S matrix among these four states. Because fermions can only be created or destroyed in pairs, the first two states scatter into each other, and so do the last two. In addition, the S matrix commutes with the operator Q_+Q_-, which in this basis is

$$Q_+Q_-=2i\begin{bmatrix} 1 & \sinh\theta & 0 & 0\\ \sinh\theta & -1 & 0 & 0\\ 0 & 0 & 0 & -\cosh\theta\\ 0 & 0 & -\cosh\theta & 0 \end{bmatrix}.$$

It follows that the S matrix is diagonal in the basis

$$\begin{aligned} |S^{ab}\rangle&=\frac{1}{[\cosh(\frac{1}{2}\theta)]^{1/2}}\left[\cosh(\tfrac{1}{4}\theta)|b^a(\tfrac{1}{2}\theta)b^b(-\tfrac{1}{2}\theta)\rangle\right.\\ &\qquad\left.+\sinh(\tfrac{1}{4}\theta)|f^a(\tfrac{1}{2}\theta)f^b(-\tfrac{1}{2}\theta)\rangle\right],\\ |T^{ab}\rangle&=\frac{1}{[\cosh(\frac{1}{2}\theta)]^{1/2}}\left[-\sinh(\tfrac{1}{4}\theta)|b^a(\tfrac{1}{2}\theta)b^b(-\tfrac{1}{2}\theta)\rangle\right.\\ &\qquad\left.+\cosh(\tfrac{1}{4}\theta)|f^a(\tfrac{1}{2}\theta)f^b(-\tfrac{1}{2}\theta)\rangle\right],\\ |U^{ab}\rangle&=\frac{1}{\sqrt{2}}\left[|b^a(\tfrac{1}{2}\theta)f^b(-\tfrac{1}{2}\theta)\rangle+|f^a(\tfrac{1}{2}\theta)b^b(-\tfrac{1}{2}\theta)\rangle\right],\\ |V^{ab}\rangle&=\frac{1}{\sqrt{2}}\left[|b^a(\tfrac{1}{2}\theta)f^b(-\tfrac{1}{2}\theta)\rangle-|f^a(\tfrac{1}{2}\theta)b^b(-\tfrac{1}{2}\theta)\rangle\right]. \end{aligned} \quad (15)$$

(Whenever we write a two-particle bra $|S^{ab}\rangle$ or ket $\langle S^{ab}|$, the first isospin index will refer to the particle of rapidity $\theta/2$ and the second will refer to the particle of rapidity $-\theta/2$.) The statement that the S matrix is diagonal in the above basis means that

$$\langle S^{ab}|S|T^{cd}\rangle=\langle U^{ab}|S|V^{cd}\rangle=0\,, \quad (16)$$

etc.

In addition, we find $Q_+|S^{ab}\rangle=[2\cosh(\frac{1}{2}\theta)]^{1/2}|U^{ab}\rangle$, $Q_+|U^{ab}\rangle=[2\cosh(\frac{1}{2}\theta)]^{1/2}|S^{ab}\rangle$, and therefore

$$0=\langle U^{cd}|[Q_+,S]|S^{ab}\rangle$$
$$=[2\cosh(\tfrac{1}{2}\theta)]^{1/2}(\langle S^{cd}|S|S^{ab}\rangle-\langle U^{cd}|S|U^{ab}\rangle)\,, \tag{17}$$

so that supersymmetry implies

$$\langle S^{cd}|S|S^{ab}\rangle=\langle U^{cd}|S|U^{ab}\rangle\,. \tag{18}$$

A similar argument shows

$$\langle T^{cd}|S|T^{ab}\rangle=\langle V^{cd}|S|V^{ab}\rangle\,. \tag{19}$$

Finally, expanding in all the possible isospin invariants, we find that the most general form for the S matrix permitted by supersymmetry and isospin is

$$\begin{aligned}\langle S^{cd}|S|S^{ab}\rangle&=\langle U^{cd}|S|U^{ab}\rangle\\&=S_1(\theta)\delta^{ac}\delta^{bd}+S_2(\theta)\delta^{ab}\delta^{cd}\\&\quad+S_3(\theta)\delta^{ad}\delta^{bc}\,,\\\langle T^{cd}|S|T^{ab}\rangle&=\langle V^{cd}|S|V^{ab}\rangle\\&=T_1(\theta)\delta^{ac}\delta^{bd}+T_2(\theta)\delta^{ab}\delta^{cd}\\&\quad+T_3\delta^{ad}\delta^{bc}\,,\end{aligned} \tag{20}$$

with the other S-matrix elements vanishing.[10]

We will also have to use the restrictions that are placed on the S matrix by crossing, unitarity, and analyticity. Crossing exchanges θ with $i\pi-\theta$. The crossing relations are simplest if written in terms of the sums and differences of the S_i and T_i. Taking account of certain factors of i that appear in exchanging a boson with a fermion under crossing, the crossing relations are

$$\begin{aligned}&S_1(i\pi-\theta)+T_1(i\pi-\theta)=S_1(\theta)+T_1(\theta)\,,\\&S_1(i\pi-\theta)-T_1(i\pi-\theta)=-i\tanh(\tfrac{1}{2}\theta)\\&\qquad\times[S_1(\theta)-T_1(\theta)]\end{aligned} \tag{21}$$

and

$$\begin{aligned}&S_2(i\pi-\theta)+T_2(i\pi-\theta)=S_3(\theta)+T_3(\theta)\,,\\&S_2(i\pi-\theta)-T_2(i\pi-\theta)=-i\tanh(\tfrac{1}{2}\theta)\\&\qquad\times[S_3(\theta)-T_3(\theta)]\,.\end{aligned} \tag{22}$$

In writing the constraints imposed by unitarity, it is convenient to decompose the S matrix into states of definite isospin. There are three isospin channels—the initial state can be symmetric and traceless, antisymmetric, or isosinglet. The amplitudes for $\langle S^{cd}|S|S^{ab}\rangle$ in the symmetric traceless, antisymmetric, and isosinglet channels are, respectively, S_1+S_3, S_1-S_3, and $S_1+S_3+NS_2$. Unitarily implies that in the physical region (real, positive θ) these amplitudes each have modulus one, so

$$\begin{aligned}&[S_1(\theta)+S_3(\theta)][S_3^*(\theta)+S_1^*(\theta)]=1\,,\\&[S_1(\theta)-S_3(\theta)][S_1^*(\theta)-S_3^*(\theta)]=1\,,\\&[S_1(\theta)+S_3(\theta)+NS_2(\theta)][S_1^*(\theta)+S_3^*(\theta)\\&\qquad+NS_2^*(\theta)]=1\,,\end{aligned} \tag{23}$$

for real θ. The amplitudes, however, are all real in the region below threshold, which corresponds to imaginary θ. It follows, using the Schwartz reflection principle, that $S_1^*(\theta)=S_i(-\theta)$ for θ real and $i=1,2,3$, so we can rewrite unitarity in the form

$$\begin{aligned}&[S_1(\theta)+S_3(\theta)][S_1(-\theta)+S_3(-\theta)]=1\,,\\&[S_1(\theta)-S_3(\theta)][S_1(\theta)-S_3(-\theta)]=1\,,\\&[S_1(\theta)+S_3(\theta)+NS_2(\theta)][S_1(-\theta)+S_3(-\theta)\\&\qquad+NS_2(-\theta)]=1\,.\end{aligned} \tag{24}$$

In this form the unitarity equations are valid throughout the complex θ plane. We have, of course, exactly analogous equations with T replacing S.

Finally, what are the analytic properties of the S_i and T_i? In terms of the usual kinematic variable $s=(p_1+p_2)^2$, we ordinarily would expect elastic thresholds at $s=4m^2$ and $s=0$, as well as inelastic thresholds at $s=16m^2$, $s=36m^2$, etc. However, in this theory, the inelastic thresholds are absent because of the absence of particle production, so the S matrix is analytic except for two-body cuts beginning at $s=4m^2$ and $s=0$. The elastic thresholds are known to be square-root branch points, and the transformation from s to θ, $s=4\cosh^2(\frac{1}{2}\theta)$, is, as noted by Zamolodchikov, exactly such as to remove these square-root singularities.[11] Therefore the S-matrix elements are expected to be meromorphic functions of θ with no cuts.

Now, of course, we must turn to the problem of determining these meromorphic S_i and T_i. We do this by using the cubic constraint on the S matrix discussed in Sec. II.

This cubic constraint, as we recall, states that a three-body collision can be regarded in two different ways as a succession of two-body collisions. For each possible three-body initial and final state, there is a separate condition that we must impose, expressing the fact that this particular three-body S-matrix element can be factorized in two ways. Since there are many possible initial and final states, the analysis is somewhat tedious.

As an example, we consider the S-matrix amplitude for

$$|b^a(\theta_0)f^b(\theta_0+\theta_1)b^c(\theta_0+\theta_1+\theta_2)\rangle$$
$$\to|b^b(\theta_0)f^c(\theta_0+\theta_1)b^a(\theta_0+\theta_1+\theta_2)\rangle\,.$$

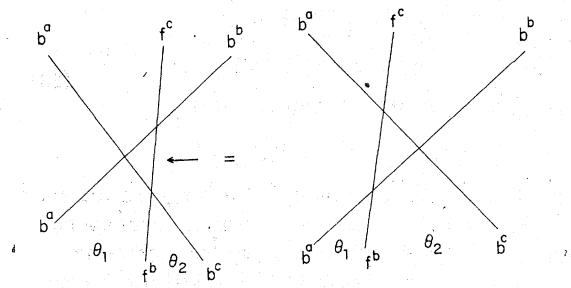

FIG. 4. Two collisions which must have the same amplitude in this theory. Equating the amplitudes gives a nontrivial constraint since each two-body collision is described by an isospin matrix.

This reaction involves isospin exchange but no spin exchange, since the outgoing particle of rapidity, say, θ_0 has different isospin from the incoming one, but has the same statistics. It will be sufficient for us, in treating this reaction, to consider only the case $a \neq b \neq c$.

As illustrated in Fig. 4, we must consider the fact that there are two ways to factorize this amplitude. (In the diagram, θ_1, and θ_2 are the rapidity differences between neighboring particles.)

In considering either sketch in Fig. 4 we must remember that a sum over all possible identities for the particles in the intermediate states is implied. For example, the line indicated by the arrow in the left half of Fig. 4 may be a boson or a fermion of any isospin. Therfore, when expanded in terms of the functions we have called S_i and T_i, either half of Fig. 4 represents a considerable number of terms. Some simplification occurs if we add Fig. 4 to the same diagram with all bosons replaced by fermions, and all fermions by bosons. The identity of the two halves of Fig. 4 then gives the requirement

$$\begin{aligned}
&[S_1(\theta_1)+T_1(\theta_1)][S_3(\theta_2)+T_3(\theta_2)][S_3(\theta_1+\theta_2)+T_3(\theta_1+\theta_2)]\\
&\quad+[S_1(\theta_1)-T_1(\theta_1)][S_3(\theta_2)-T_3(\theta_2)][S_3(\theta_1+\theta_2)-T_3(\theta_1+\theta_2)]\\
&\quad+[S_3(\theta_1)+T_3(\theta_1)][S_1(\theta_2)+T_1(\theta_2)][S_3(\theta_1+\theta_2)+T_3(\theta_1+\theta_2)]\\
&\quad+[S_3(\theta_1)-T_3(\theta_1)][S_1(\theta_2)+T_1(\theta_2)][S_3(\theta_1+\theta_2)+T_3(\theta_1+\theta_2)]\\
&\qquad=[S_3(\theta_1)+T_3(\theta_1)][S_3(\theta_2)+T_3(\theta_2)][S_1(\theta_1)+T_1(\theta_1)]\\
&\qquad\quad+[S_3(\theta_1)-T_3(\theta_1)][S_3(\theta_2)-T_3(\theta_2)][S_1(\theta_1+\theta_2)-T(\theta_1+\theta_2)]\,,
\end{aligned}\tag{25}$$

where the left- and right-hand sides of the equation correspond to the left- and right-hand sides of the figure.

If we consider instead the reaction

$$|f^a(\theta_0)b^b(\theta_0+\theta_1)b^c(\theta_0+\theta_1+\theta_2)\rangle \to |f^b(\theta_0)f^c(\theta_0+\theta_1)b^a(\theta_0+\theta_1+\theta_2)\rangle$$

(and the same reaction with bosons and fermions interchanged), we obtain a similar equation, but with a different coefficient for the second, fourth, and sixth terms. These equations combine to give

$$\begin{aligned}
&[S_1(\theta_1)+T_1(\theta_1)][S_3(\theta_2)+T_3(\theta_2)][S_3(\theta_1+\theta_2)+T_3(\theta_1+\theta_2)]\\
&\quad+[S_3(\theta_1)+T_3(\theta_1)][S_1(\theta_2)+T_1(\theta_2)][S_3(\theta_1+\theta_2)+T_3(\theta_1+\theta_2)]\\
&\qquad=[S_3(\theta_1)+T_3(\theta_1)][S_3(\theta_2)+T_3(\theta_2)][S_1(\theta_1+\theta_2)+T_1(\theta_1+\theta_2)]\,,
\end{aligned}\tag{26}$$

and

$$\begin{aligned}
&[S_1(\theta_1)-T_1(\theta_1)][S_3(\theta_2)-T_3(\theta_2)][S_3(\theta_1+\theta_2)-T_3(\theta_1+\theta_2)]\\
&\quad+[S_3(\theta_1)-T_3(\theta_1)][S_1(\theta_2)-T_1(\theta_2)][S_3(\theta_1+\theta_2)-T_3(\theta_1+\theta_2)]\\
&\qquad=[S_3(\theta_1)-T_3(\theta_1)][S_3(\theta_2)-T_3(\theta_2)][S_1(\theta_1+\theta_2)-T_1(\theta_1+\theta_2)]\,.
\end{aligned}\tag{27}$$

The meaning of the first equation can be clarified by dividing by

$$[S_3(\theta_1)+T_3(\theta_1)][S_3(\theta_2)+T_3(\theta_2)][S_3(\theta_1+\theta_2)+T_3(\theta_1+\theta_2)]\,.$$

We find

$$\frac{S_1(\theta_1)+T_1(\theta_1)}{S_3(\theta_1)+T_3(\theta_1)}+\frac{S_1(\theta_2)+T_1(\theta_2)}{S_3(\theta_2)+T_3(\theta_2)}=\frac{S_1(\theta_1+\theta_2)+T_1(\theta_1+\theta_2)}{S_3(\theta_1+\theta_2)+T_3(\theta_1+\theta_2)}\,.\tag{28}$$

In other words, the function $(S_1+T_1)/(S_3+T_3)$ is a linear function of θ. So

$$\frac{S_1(\theta)+T_1(\theta)}{S_3(\theta)+T_3(\theta)}=i\lambda\theta \tag{29}$$

for some constant λ, which we will soon determine. (Unitarity will require that λ is real.) Likewise

$$\frac{S_1(\theta^1)-T_1(\theta^1)}{S_3(\theta^1)-T_3(\theta^1)}=i\rho\theta \tag{30}$$

for some real constant ρ.

If we now consider the amplitude corresponding to Fig. 4 and *subtract* from it the amplitude for the reaction with bosons and fermions exchanged, the resulting equation implies $\lambda=\rho$.

Finally, it is convenient to consider the reaction

$$|f^a(\theta_0)b^b(\theta_0+\theta_1)b^c(\theta_0+\theta_1+\theta_2)\rangle \to |b^b(\theta_0)b^c(\theta_0+\theta_1)f^a(\theta_0+\theta_1+\theta_2)\rangle .$$

If one imposes the cubic condition on this S-matrix element, one learns that

$$\frac{S_1(\theta)-T_1(\theta)}{S_1(\theta)+T_1(\theta)}=\frac{if}{\sinh(\frac{1}{2}\theta)} \tag{31}$$

for some constant f.

It turns out that there is no need to consider the cubic constraint on the S matrix any further, since the equations corresponding to the other initial and final three-body states either are identities, or are consequences of the relations that we will now deduce more easily from crossing symmetry and unitarity.

The crossing equations have a dramatic consequence—as in the case considered by Zamolodchikov, they will determine the functions S_2 and T_2. In fact, combining (21), (22), and (29) yields

$$\begin{aligned} S_2(\theta)+T_2(\theta)&=S_3(i\pi-\theta)+T_3(i\pi-\theta)\\ &=\frac{S_1(i\pi-\theta)+T_1(i\pi-\theta)}{i\lambda(i\pi-\theta)}\\ &=\frac{S_1(\theta)+T_1(\theta)}{i\lambda(i\pi-\theta)}, \end{aligned} \tag{32}$$

and combining (21), (22), and (30) yields

$$\begin{aligned} S_2(\theta)-T_2(\theta)&=\frac{i}{\tanh(\frac{1}{2}\theta)}[S_3(i\pi-\theta)-T_3(i\pi-\theta)]\\ &=\frac{i}{\tanh(\frac{1}{2}\theta)}\left[\frac{S_1(i\pi-\theta)-T_1(i\pi-\theta)}{i\lambda(i\pi-\theta)}\right]\\ &=\frac{S_1(\theta)-T_1(\theta)}{i\lambda(i\pi-\theta)}, \end{aligned} \tag{33}$$

and adding and subtracting these equations, we find

$$\begin{aligned} S_2(\theta)&=\frac{S_1(\theta)}{i\lambda(i\pi-\theta)},\\ T_2(\theta)&=\frac{T_1(\theta)}{i\lambda(i\pi-\theta)}. \end{aligned} \tag{34}$$

Now, Eq. (31) implies that we can write

$$\begin{aligned} S_1(\theta)&=\left(1-\frac{if}{\sinh(\frac{1}{2}\theta)}\right)S(\theta),\\ T_1(\theta)&=\left(1+\frac{if}{\sinh(\frac{1}{2}\theta)}\right)S(\theta), \end{aligned} \tag{35}$$

where $S(\theta)$ is a new unknown function. In view of (29), (30), (34), and (35), S_1, S_2, S_3, T_1, T_2, and T_3 are all known in terms of S, λ, and f. If we substitute this information into the unitarity equations, we learn that $\lambda=-(N-2)/2\pi$ and

$$S(\theta)S(-\theta)=\frac{\theta^2}{\theta^2+\Delta^2}\,\frac{\sinh^2(\frac{1}{2}\theta)}{\sinh^2(\frac{1}{2}\theta)+f^2}, \tag{36}$$

where we have let $\Delta=-1/\lambda=2\pi/(N-2)$. Crossing symmetry, moreover, now implies

$$S(\theta)=S(i\pi-\theta). \tag{37}$$

To summarize our results, we now have

$$\begin{aligned} S_1(\theta)&=\left(1-\frac{if}{\sinh(\frac{1}{2}\theta)}\right)S(\theta),\\ S_2(\theta)&=\frac{2\pi i}{N-2}\,\frac{S_1(\theta)}{(i\pi-\theta)},\\ S_3(\theta)&=\frac{2\pi i}{N-2}\,\frac{S_1(\theta)}{\theta},\\ T_1(\theta)&=\left(1+\frac{if}{\sinh(\frac{1}{2}\theta)}\right)S(\theta),\\ T_2(\theta)&=\frac{2\pi i}{N-2}\,\frac{T_1(\theta)}{i\pi-\theta},\\ T_3(\theta)&=\frac{2\pi i}{N-2}\,\frac{T_1(\theta)}{\theta}, \end{aligned} \tag{38}$$

where f is an unknown constant, and $S(\theta)$ is an unknown meromorphic function that satisfies (36) and (37).

To determine f requires some information of a different sort. At least for large enough N,[4] it is known that the interaction is repulsive in the channels corresponding to S_1, S_2, and S_3 but attractive in the channels T_1, T_2, and T_3. Therefore we expect bound-state poles in the T_i but not the S_i. But a bound-state pole present in the T_i can be absent in the S_i only if it is canceled in the S_i by the factor $[1-if/\sinh(\frac{1}{2}\theta)]$. Therefore we expect that the function $S(\theta)$ contains only a single bound-state pole at $\theta=\theta_0$ where $\sinh(\frac{1}{2}\theta_0)=if$.

Furthermore, the large-N expansion indicates that bound-state poles will be present in T_2 and in T_1-T_3 but not in T_1+T_3. (This can be easily

understood. For large N, the interaction is weak, and the bound state is near threshold, where the T_i describe almost pure fermion-fermion scattering. The T_1+T_3 amplitude is symmetric in isospin, so it corresponds to an antisymmetric wave function in coordinate space, and a weak attraction will not produce a bound state in such a channel.) How can a bound-state pole present in T_1 and in T_1-T_3 be absent in T_1+T_3? This is possible only if $1+2\pi i/(N-2)\theta_0=0$, where θ_0 is the position of the pole. So $\theta_0=-2\pi i/(N-2)$ and, combining this with our previous result, we find that $f=\sin[\pi/(N-2)]$.

Now that f has been determined, all that remains is to determine the meromorphic function $S(\theta)$ which satisfies

$$S(\theta)S(-\theta)=\frac{\theta^2}{\theta^2+\Delta^2}\,\frac{\sinh^2(\frac{1}{2}\theta)}{\sinh^2(\frac{1}{2}\theta)+\sin^2(\frac{1}{2}\Delta)},$$

and $S(\theta)=S(i\pi-\theta)$ and which has a single bound-state pole at $\theta=i\Delta$ and no other poles in the physical region ($0\leq \mathrm{Im}\,\theta\leq\pi$). To construct such a function, we first write $S(\theta)=Q(\theta)Y(\theta)$ where Q and Y are to satisfy

$$Q(\theta)Q(-\theta)=\frac{\theta^2}{\theta^2+\Delta^2},\qquad Y(\theta)Y(-\theta)=\frac{\sinh^2(\frac{1}{2}\theta)}{\sinh^2(\frac{1}{2}\theta)+\sin^2(\frac{1}{2}\Delta)},\tag{39}$$

and also

$$Q(\theta)=Q(i\pi-\theta),\qquad Y(\theta)=Y(i\pi-\theta).\tag{40}$$

Let us first discuss the construction of $Q(\theta)$. If we only had to satisfy (39), a simple choice would be

$$Q(\theta)=\frac{\theta}{\theta+i\Delta}.\tag{41}$$

This violates the crossing relation (40) which we might repair by writing

$$Q(\theta)=\frac{\theta}{\theta+i\Delta}\,\frac{i\pi-\theta}{i\pi-\theta+i\Delta}.\tag{42}$$

Now (39) is no longer satisfied, a situation which can be redressed by writing

$$Q(\theta)=\frac{\theta}{\theta+i\Delta}\,\frac{i\pi-\theta}{i\pi-\theta+i\Delta}\,\frac{i\pi+\theta+i\Delta}{i\pi+\theta}.\tag{43}$$

Now we have ruined crossing symmetry again. It can be restored by including an extra factor:

$$Q(\theta)=\frac{\theta}{\theta+i\Delta}\,\frac{i\pi-\theta}{i\pi-\theta+i\Delta}\,\frac{i\pi+\theta+i\Delta}{i\pi+\theta}\times\frac{2i\pi-\theta+i\Delta}{2i\pi-\theta}.\tag{44}$$

Again (39) is ruined. Continuing in this way, we obtain an infinite product which converges to the desired function $Q(\theta)$. The resulting $Q(\theta)$ can be rewritten as a product of Γ functions $Q(\theta)=R(\theta)R(i\pi-\theta)$, where

$$R(\theta)=\frac{\Gamma(\Delta/2\pi-i\theta/2\pi)\Gamma(\frac{1}{2}-i\theta/2\pi)}{\Gamma(-i\theta/2\pi)\Gamma(\frac{1}{2}+\Delta/2\pi-i\theta/2\pi)}.\tag{45}$$

A similar iteration for $Y(\theta)$, starting from the first guess

$$Y(\theta)=\frac{\sinh(\frac{1}{2}\theta)}{\sinh(\frac{1}{2}\theta)-i\sin(\frac{1}{2}\Delta)}\tag{46}$$

leads to

$$Y(\theta)=\frac{\sinh(\frac{1}{2}\theta)}{\sinh(\frac{1}{2}\theta)-i\sin(\frac{1}{2}\Delta)}\,\frac{\sinh[\frac{1}{2}(i\pi-\theta)]}{\sinh[\frac{1}{2}(i\pi-\theta)-i\sin(\frac{1}{2}\Delta)]}\times\frac{\sinh[\frac{1}{2}(i\pi+\theta)]+i\sin(\frac{1}{2}\Delta)}{\sinh[\frac{1}{2}(i\pi+\theta)]}\,\frac{\sinh[\frac{1}{2}(2i\pi-\theta)]+i\sin(\frac{1}{2}\Delta)}{\sinh[\frac{1}{2}(2i\pi-\theta)]}\cdots.\tag{47}$$

This can also be written as an infinite product of Γ functions, $Y(\theta)=R(\theta)R(i\pi-\theta)$, where

$$R(\theta)=\left[\frac{\Gamma(-i\theta/2\pi)}{\Gamma(\frac{1}{2}-i\theta/2\pi)}\prod_{l=1}^{\infty}\frac{\Gamma(\Delta/2\pi-i\theta/2\pi+l)\Gamma(-i\theta/2\pi-\Delta/2\pi+l-1)\Gamma^2(-i\theta/2\pi+l-\frac{1}{2})}{\Gamma(\Delta/2\pi-i\theta/2\pi+l+\frac{1}{2})\Gamma(-i\theta/2\pi-\Delta/2\pi+l-\frac{1}{2})\Gamma^2(-i\theta/2\pi+l-1)}\right].\tag{48}$$

We believe that the solution $S(\theta)=Q(\theta)Y(\theta)$, with Q and Y as given above, gives the exact S matrix for the supersymmetric nonlinear σ model. Given our physical requirement that there is a bound-state pole at $\theta=i\Delta$ and no others, it is the unique solution with no Castillejo-Dalitz-Dyson (CDD) zeros in the physical region. It impresses us as being the simplest solution. And it agrees with the results of Alvarez for the large-N expansion. This last point is a very important check on the calculation.

We have factored the S matrix as a product of two functions, Q and Y. What is the meaning of this factorization?

$Q(\theta)$ is the function encountered by Zamolodchikov in solving the nonlinear σ model. We claim

that $Y(\theta)$ would appear in the same way in solving the supersymmetric sine-Gordon model, which is described by the Lagrangian

$$\mathcal{L}=\tfrac{1}{2}(\partial_\mu\phi)^2+\tfrac{1}{2}i\bar{\psi}\partial\!\!\!/\psi+\frac{1}{4\beta^2}\cos^2\beta\phi-\tfrac{1}{2}(\cos\beta\phi)\bar{\psi}\psi\,, \tag{49}$$

where ϕ is a real scalar field, ψ is a Majorana fermion, and β is a coupling constant. If we consider in this model the scattering of the elementary particles—bosons and fermions—then the supersymmetry constraints leave two invariant amplitudes, which can be determined by the methods we have discussed and turn out to be

$$\left\{1\pm i\frac{\sin(\frac{1}{2}\Delta)}{\sinh(\frac{1}{2}\theta)}\right\}Y(\theta)\,,$$

with the function $Y(\theta)$ described above. (One must make a correct identification of Δ in terms of β; this can be done by studying the high-energy limit of the S matrix.) The S matrix of the supersymmetric nonlinear σ model seems to be a product of two S matrices—an isospin-bearing S matrix, which is the S matrix of the ordinary nonlinear σ model, and a Bose-Fermi S matrix, which is the S matrix of the supersymmetric sine-Gordon equation.

The two-body S matrix that we have constructed contains, as we have said, a single bound-state pole at $\theta=i\Delta$, which appears in the channels T_1-T_2 and T_3. For large N, these are essentially fermion-fermion bound states.

If one considers the multiparticle S-matrix elements—which are known from factorization from the two-particle S-matrix elements—then, as in the case treated in Ref. 7, one finds additional multifermion bound states. In fact, the n-body scattering amplitude contains a bound-state pole with a mass proportional to $\sin[n\pi/(N-2)]$. This coincides with the mass spectrum found by Dashen, Hasslacher, and Neveu[12] and by Zamolodchikov[7] for the Gross-Neveu model. In addition, the isospin quantum numbers of the fermion bound states are the same in the model we are treating as in the Gross-Neveu model—for instance, the two-body bound states appear in the isosinglet channel (T_2) and the antisymmetric tensor channel (T_1-T_3).

This is certainly a striking similarity between the supersymmetric σ model and the Gross-Neveu model. It tends to confirm the idea of Alvarez[4] that the supersymmetric σ model is a kind of amalgam of the ordinary σ model and the Gross-Neveu model, equally suspended between the two. Alvarez shows, for example, that to lowest order in $1/N$, the boson-boson scattering amplitudes of the supersymmetric σ model coincide with those of the ordinary σ model, while the fermion-fermion scattering amplitudes of the supersymmetric σ model coincide with those of the Gross-Neveu model.

There is also a more mathematical sense in which our results confirm this idea.

Let us consider the process that we followed in constructing Q and Y. In constructing Q by iteration, we started with a first quess

$$Q=\frac{\theta}{\theta+i\Delta}\,. \tag{50}$$

But might we not equally well have started with a first guess

$$Q=\frac{\theta}{\theta-i\Delta}\,? \tag{51}$$

In constructing Y, we started with

$$Y(\theta)=\frac{\sinh(\frac{1}{2}\theta)}{\sinh(\frac{1}{2}\theta)-i\sin(\frac{1}{2}\Delta)}\,. \tag{52}$$

Why did we not start with

$$Y(\theta)=\frac{\sinh(\frac{1}{2}\theta)}{\sinh(\frac{1}{2}\theta)+i\sin(\frac{1}{2}\Delta)}\,? \tag{53}$$

At first sight, we seem, in fact, to have four acceptable starting points. We may take (50) or (51) and (52) or (53). However, we want our S matrix to have a single bound-state pole at $\theta=i\Delta$. If we start with (50) and (53), we obtain an S matrix with no bound-state pole, while starting from (51) and (52) we get an S matrix with a double pole at $\theta=i\Delta$. Both are unacceptable.

But we have two acceptable starting points—(50) with (52) or (51) with (53). Either leads to a single pole at $\theta=i\Delta$. So we seem to have two possible, and equally natural, S matrices.

In fact, in the problems considered by Zamolodchikov and Zamolodchikov, they were led to construct only the function we have called Q (Y did not appear). There are two natural starting points, (50) and (51), and therefore two S matrices. One turns out to be the S matrix of the nonlinear σ model, while the other is the S matrix of the Gross-Neveu model.

Likewise, in our case, we might expect to have two S matrices, one for the supersymmetric σ model and one for a supersymmetric Gross-Neveu model.

But surprisingly, the iterative procedure gives the same result for the S matrix (which depends only on the product $S=QY$) whether one starts with (50) and (52) or with (51) and (53). (This results from some algebraic identities for the Γ functions.) Yet (51) and (53) are the starting point closest to the Gross-Neveu model, while (50) and (52) are the starting point closest to the nonlinear σ model.

Our S matrix thus corresponds to the supersymmetric versions of both the Gross-Neveu model and the nonlinear σ model.

IV. CONCLUSION

Now that we have a proposed S matrix for the model, we would like to return to the physical questions which motivated this research.

We wanted first to know whether the dynamical breaking of the γ_5 symmetry and the mass generation which appear in the $1/N$ expansion are properties of the exact theory. Our results seem to confirm that these properties are exact. The proposed S matrix agrees with the $1/N$ expansion and with all the general properties of quantum field theory (unitarity, analyticity, crossing). If this S matrix is correct, it certainly shows that the model contains symmetry breaking and mass generation.

Because of the breaking of the discrete γ_5 symmetry, the full spectrum of this theory will contain "kink" states, similar to those found in a semiclassical calculation by Callan, Coleman, Gross, and Zee[13] for the $(\bar{\psi}\psi)^2$ model, in addition to the elementary particles and their bound states. We have, of course, found only the S-matrix elements for the elementary particles (from which the bound-state amplitudes can be determined by factorization). The S matrix for the kinks remains unknown.

We also wanted to know if this theory exhibits some special behavior at $N=3$, which is the case most closely related to four-dimensional gauge theories. Here the results are less conclusive.

Although well behaved for $N>4$, the S matrix that we have constructed is anomalous in a variety of ways for $N\leq 4$. One of the bound-state masses goes to zero as $N\to 4$ from above; the threshold behavior of the S matrix, although correct for $N>3$, is wrong for $N=3$; and the S matrix we have computed for $N=3$ lacks the enlarged supersymmetry algebra that the theory is known to possess for this value of N.[2]

A resolution of these problems (and of similar questions about the Gross-Neveu model) has been suggested by A. B. Zamolodchikov (private communication). By way of motivation, let us recall that in the sine-Gordon theory, for $\beta^2>4\pi$, the elementary boson disappears from the spectrum, which consists only of kinks. Motivated in part by semiclassical mass formulas, Zamolodchikov suggests that a similar phenomenon occurs in these models for $N\leq 4$: the elementary particles disappear from the spectrum, and only the kink states survive. Thus, to understand the theory for $N=3$ or $N=4$, it would be necessary to determine the kink S matrix.

Finally, we are left wondering whether or not the two-dimensional nonlinear σ model has an important lesson to teach about four-dimensional theories. Could there be an S-matrix method for dealing with four-dimensional quantum chromodynamics? That would presumably require the existence of some hitherto unguessed symmetry structure (which, because, of the Coleman-Mandula theorem, could not consist of the existence of conserved local charges). Or would an understanding of how the σ model S matrix comes about from a "microscopic" point of view help us in understanding the structure of four-dimensional chromodynamics?

ACKNOWLEDGMENTS

We would like to thank O. Alvarez, S. Coleman, D. Kayhdan, and M. Peskin for discussions. One of us (R. Shankar) would like to thank the Harvard Society of Fellows for its support while part of this work was being carried out. The work of E. Witten was supported in part by the Harvard Society of Fellows.

This research was supported in part by the National Science Foundation under Grant No. PHY75-20427.

[1]P. DiVecchia and S. Ferrara, Nucl. Phys. B130, 93 (1977).

[2]E. Witten, Phys. Rev. D 16, 2991 (1977).

[3]S. Libby has made the interesting suggestion (private communication) of instead coupling fermions to the σ model with a term $\bar{\psi}_i\gamma_\mu\psi_j(n_i\partial_\mu n_j - n_j\partial_\mu n_i)$.

[4]O. Alvarez, Phys. Rev. D 17, 1123 (1978).

[5]A. B. Zamolodchikov, Zh. Éksp. Teor. Fiz. Pis'ma Red. 25, 499 (1977) [JETP Lett. 25, 468 (1977)]; Report No. ITEP-12, 1977 (unpublished).

[6]M. Karowski, H. J. Thun, T. T. Truong, and P. H. Weisz, Phys. Lett. 67B, 321 (1977).

[7]A. Zamolodchikov and Al. Zamolodchikov, Report No. JINRE-210857, 1977 (unpublished); Report No. ITEP-112 (unpublished).

[8]Although some of the presentation is new, this section is largely a review of work by A. M. Polyakov (unpublished); and R. Flume, V. Glaser, and D. Jagolnitzer (unpublished); and P. P. Kulish, Teor. Mat. Fiz. 26, 198 (1976). The first use of these methods to construct an S matrix was Ref. 6.

[9]S. Coleman and J. Mandula, Phys. Rev. 159, 1251 (1967).

[10]We have defined the S matrix with the δ function for momentum conservation removed.

[11]This change of variables also eliminates a logarithmic branch point at infinity.

[12]R. Dashen, B. Hasslacher, and A. Neveu, Phys. Rev. D 12, 2443 (1975).

[13]C. G. Callan, Jr., S. Coleman, D. J. Gross, and A. Zee (unpublished); see Ref. 12 for a discussion of this work.