PHYSICAL REVIEW D VOLUME 18, NUMBER 6 15 SEPTEMBER 1978

Grand partition function in field theory with applications to sine-Gordon field theory

Stu Samuel
Department of Physics, University of California, Berkeley, California 94720
(Received 24 February 1978)

Certain relativistic field theories are shown to be equivalent to the grand partition function of an interacting gas. Using the physical insight given by this analogy many field-theoretic results are obtained, particularly for the sine-Gordon field theory. The main results are enumerated in the summary to which the reader is referred.

I. INTRODUCTION

This paper will employ a technique, known as Gaussian integration,[1] by which certain field theories are identified with a gas of interacting particles. Originally the purpose was to rewrite a partition function in a field-theoretic way, so that field-theory techniques could be used.[2] The idea of this paper is to reverse the process in cases where the analog field theory is a relativistic one. An example is the sine-Gordon field theory which is equivalent to the theory of a neutral Coulomb gas.[3] The vacuum expectation value of $e^{\int \mathcal{L}_I} = e^{2\lambda_0 \int \cos \beta q \chi}$, which is a sum of vacuum bubble diagrams, is equal to the grand partition function for such a Coulomb system, λ_0 playing the role of the absolute activity and β playing the role of the inverse temperature. In these cases one can analyze the field theory by using the underlying statistical-mechanical analog. Knowledge of the sine-Gordon theory will yield information about the Coulomb plasma. Likewise, one may use the Coulomb plasma to gain information about the sine-Gordon theory. This is the plan of this paper. It enables one to use one's intuition of the Coulomb plasma to obtain field-theoretic results.

Some of the results of this paper have appeared in the mathematics literature.[3-5] The author feels these are worth repeating since such mathematical presentations are not accessible to most physicists. This paper stresses simple, physical, and intuitive methods of derivation.

The paper is organized as follows: Section II reviews the Gaussian representation method and Sec. III is a perturbative check. This check gives one insight into the statistical-mechanical–field-theoretic analogy. In particular, the Feynman diagrams have a simple physical description in terms of the underlying thermodynamic system. This correspondence is outlined in Table I. From Sec. IV onward the main concern of the paper is the two-dimensional sine-Gordon theory. Section IV introduces the sine-Gordon field theory and discusses its infrared-singular nature which in the Coulomb analog model forces strict neutrality. Section V determines the phases of the sine-Gordon theory. At low temperatures there is a dipole gas, whereas at high temperatures there is a plasma phase. The impact on the existence of solitons is discussed. Section VI shows how the nonlinear σ model is equivalent to the sine-Gordon theory. In Sec. VII the renormalization is performed to all orders in λ_R and βq^2, when βq^2 is small. This shows that the theory is well defined. Other aspects of renormalization are also dealt with. In Sec. VIII remaining ideas are discussed, most of which depend heavily on the Coulomb gas analogy. Most important is the vacuum structure and its effect on the theory. Charge screening and shielding, fractional charges, the effects of infrared divergences on Feynman rules, and the sine-Gordon solitions are discussed. Section IX is the summary. There, the main results are simply enumerated. The paper concludes with a comment on vacuum gases as models for hadrons.

II. GAUSSIAN REPRESENTATION

The partition function for a system of interacting particles may be represented as a field theory. The technique, known as Gaussian representation, is well known in statistical mechanics.[1,2] In certain cases, the resulting field theory is a relativistic field theory. This is of particular interest since it allows one to think in terms of the underlying statistical-mechanical system. This association provides physical insight into the field theory which one usually does not have, thus allowing for the extraction of the interesting physical effects.

This section will review the Gaussian representation method along with comments concerning the use of different potentials, the smearing of fields, and various other technicalities which occur in passing from the grand partition function to the field theory. For simplicity the Gaussian representation method will be first applied to a specific example: the Coulomb plasma in three dimensions. This was actually done by Polyakov[6] in analyzing a three-dimensional instanton confinement mechanism. His instantons were monopoles interacting via a potential which was of the Coulomb type for large distances and mitigated for short distances. Because of the softened short-distance behavior, the Polyakov model has a natural renormalization prescription. This will be obvious later on. For the true Coulomb gas there is no natural renormalization and the grand partition function will be ultraviolet singular. For the present, ignore the bad short-distant behavior and any infinities which result from the use of the bare $1/r$ potential.

Consider the grand partition function $\mathfrak{z}$ for a plasma containing an arbitrary number of positive charges (ions) and negative charges (electrons) interacting via a Coulomb potential at a temperature $1/\beta$ and having an absolute activity, λ_0:

$$\mathfrak{z} = \sum_{p=0}^{\infty} \frac{\lambda_0^p}{p!} \sum_{i=0}^{\infty} \frac{\lambda_0^i}{i!} \int_V d^3R_1 \cdots d^3R_p \int_V d^3X_1 \cdots d^3X_i \exp\left[-\frac{\beta q^2}{2}\left(\sum_{l\neq m}^{p} \frac{1}{|\vec{R}_l - \vec{R}_m|} + \sum_{l\neq m}^{i} \frac{1}{|\vec{X}_l - \vec{X}_m|} - \sum_{l=1}^{p}\sum_{m=1}^{i} \frac{1}{|\vec{R}_l - \vec{X}_m|}\right)\right]. \quad (2.1)$$

In Eq. (2.1) the charge is q on both the ion and the electron. Both species have the same activity so that although the system need not be neutral, only configurations which are nearly neutral should contribute to $\mathfrak{z}$ as we know from physical considerations. λ_0, since it is the activity, is related to the chemical potential μ_0 by $\lambda_0 = e^{\beta\mu_0}$. In Eq. (2.1) V is the volume of interest, i.e., the charges are confined to the region V. Of course, $\mathfrak{z}$ does not exist because of the infinity resulting when a plus charge approaches a minus charge. One can introduce repulsive cores; alternatively one can smear the charges a bit. The latter procedure is more natural since, as will be shown later, the smearing of charges corresponds to the smearing of fields, a practice which naturally occurs in the rigorous mathematical treatment of field theories.

Except for a self-energy term, $(q^2n/2)(1/|\vec{0}|)$ which is infinite, the exponent in Eq. (2.1) may be written as

$$\frac{\iint \mathfrak{D}\chi \exp\left[-\beta\int_{R^3} \frac{\nabla\chi\cdot\nabla\chi}{8\pi} d^3R + i\beta\int_V \rho(R)\chi(R)d^3R\right]}{\iint \mathfrak{D}\chi \exp\left(-\beta\int_{R^3}\frac{\nabla\chi\cdot\nabla\chi}{8\pi}d^3R\right)}, \quad (2.2)$$

with the "charge density"

$$\rho(\vec{R};\vec{R}_1,\ldots,\vec{R}_n;q_1,\ldots,q_n) = \sum_{l=1}^{n} q_l\delta^3(\vec{R}-\vec{R}_l). \quad (2.3)$$

The self-energy infinity is made finite by smearing the charges (which is equivalent to smearing fields) or is completely eliminated by normal-ordering the final Lagrangian of Eq. (2.4) as is revealed in perturbation theory.

Doing the summations in Eq. (2.1) yields

$$\mathfrak{z} = \frac{1}{N}\iint \mathfrak{D}\chi \exp\left\{-\beta\int_{R^3}\frac{\nabla\chi\cdot\nabla\chi}{8\pi} + 2\lambda_0\int_V \cos[\beta q\chi(R)]d^3R\right\}, \quad (2.4)$$

where

$$N = \iint \mathfrak{D}\chi \exp\left(-\beta\int_{R^3}\frac{\nabla\chi\cdot\nabla\chi}{8\pi}\right).$$

Equation (2.4) is the fundamental Gaussian representation of the grand canonical sum for a Coulomb plasma. $\mathfrak{z}$ has been expressed as the field-theoretic

$$\left\langle \exp\left(2\lambda_0\int\cos\beta q\chi\right)\right\rangle,$$

where the angular brackets represent an average with respect to the free massless Euclidean functional measure in three dimensions. The corresponding Lagrangian is the sine-Gordon theory and hence one has the result that *the sine-Gordon field theory is equivalent to the Coulomb plasma.*

χ, in some sense, represents a coarse-grained Coulomb potential. The equation of motion for χ is

$$\nabla^2\chi = 4\pi(2\lambda_0 q)\sin\beta q\chi. \quad (2.5)$$

Let $\phi = i\chi$. ϕ then satisfies

$$\nabla^2\phi = -4\pi(\lambda_0 q)(-2\sinh\beta q\phi), \quad (2.6)$$

which is the well-known Debye-Hückel equation.[7] The Debye-Hückel equation is usually derived by assuming that the Coulomb potential ϕ satisfies $\nabla^2\phi(x) = -4\pi\rho(x)$ where $\rho(x)$ is the local charge density and hence equal to a mean charge density n_0 times the Boltzmann factor for a plus charge to be at x ($\exp[-\beta q\phi(x)]$) minus the Boltzmann factor for a minus charge to be at x ($\exp[\beta q\phi(x)]$). For high temperatures $\lambda_0 q = n_0$ so that the Debye-Hückel derivation yields the same result as Gaussian integration. The Debye-Hückel derivation is, at best, heuristic. For example, it is not clear why one should use the Boltzmann factors, $\exp[\pm\beta q\phi(x)]$, rather than the probability factors,

$$\exp[\pm \beta q\phi(x)]/\{\exp[\beta q\phi(x)]+\exp[-\beta q\phi(x)]\}.$$

Gaussian integration eliminates this guesswork. It tells us that the correct charge density factor is $\lambda_0 q$ when the Boltzmann factor $\exp[-\beta q\phi(x)] - \exp[\beta q\phi(x)]$ is used.

In the above example the equation of motion of the field theory corresponds to the Debye-Hückel equation of the Coulomb plasma. The Gaussian representation method applies to systems interacting via arbitrary two-body potentials. *Using the field-theory representations of these statistical-mechanical systems, one can obtain the analogy of the Debye-Huckel equation by looking at the corresponding field-theory equations of motion.*

The previous derivation may be generalized in several different ways. First of all, it does not depend on the dimension. One merely replaces the integrals in the action by d-dimensional integrals.

Secondly, one can use other potentials such as the Yukawa $\exp(-mr)/r$. Consider a potential $V(r,r')$. Let H_0 be the inverse of V, so that $H_0(r,r')\equiv\langle r|H_0|r'\rangle$ satisfies

$$\int f(r)H_0(r,r')\,V(r',r'')\,g(r'')\,d^3r\,d^3r'd^3r'' = \int f(r)\,g(r)\,d^3r$$

(for reasonable f and g). One needs to assume that

$$\int f(r)H_0(r,r')f(r')\,d^3r\,d^3r' \geq 0$$

for all reasonable f. The partition function for particles interacting via V with activity and inverse temperature respectively λ_0 and β is

$$\mathfrak{z} = \sum_n \frac{\lambda_0^{\,n}}{n!}\int d^3R_1\cdots d^3R_n \times\exp\left[-\frac{\beta}{2}\sum_{i\neq j}V(R_i,R_j)\right], \qquad (2.7)$$

which may be expressed in terms of functional integrals using the Gaussian integration method:

$$\mathfrak{z} = \frac{1}{N}\iint \mathfrak{D}\chi\exp\left[-\frac{\beta}{2}\int d^3r\,d^3r'\chi(r)H_0(r,r')\chi(r')\right] \times\exp\left(\lambda_0\int_V e^{i\chi(r)\beta}d^3r\right), \qquad (2.8)$$

where

$$N=\iint \mathfrak{D}\chi\exp\left[-\frac{\beta}{2}\int d^3r\,d^3r'\chi(r)H_0(r,r')\chi(r')\right].$$

Lastly, one can have a gas of several particles with different charges $q^{(1)}, q^{(2)},\ldots,q^{(m)}$ and activities $\lambda_0^{(1)},\ldots,\lambda_0^{(m)}$. In the Yukawa case one would call the q's quanta rather than charges. The two-body potential would be q_1q_2/r in the case of the Coulomb gas, $q_1q_2e^{-mr}/r$ in the case of a Yukawa gas, and $q_1q_2V(r)$ in the case of a general gas where $V(r)$ represents the basic potential between two positive unit quanta. The grand canonical sum is

$$\mathfrak{z} = \frac{1}{N}\iint \mathfrak{D}\chi\exp\left[-\frac{\beta}{2}(\chi,H_0\chi) + \int_V\sum_{i=1}^{m}\lambda_0^{(i)}e^{i\beta q^{(i)}\chi(r)}d^3r\right]. \qquad (2.9)$$

In particular, when $m=2$, $\lambda_0^{(1)}=\lambda_0^{(2)}=\lambda_0$, $q^{(1)}=-q^{(2)}=q$, and the Coulomb potential is used, Eq. (2.9) reduces to Eq. (2.4). By choosing $\lambda_0^{(1)}=\lambda_0^{+}$ different from $\lambda_0^{(2)}=\lambda_0^{-}$ one can deal with a Coulomb plasma of ions and electrons with an excess of ions (or electrons). Finally, for a neutral system of Yukawa particles with quanta $+q$ and $-q$ one obtains a field theory whose underlying Lagrangian is the "massive" sine-Gordon theory,

$$\mathfrak{L}=\beta\,\frac{\nabla\chi\cdot\nabla\chi}{8\pi}+\beta\,\frac{m^2\chi^2}{8\pi}-2\lambda_0\cos\beta q\chi\,.$$

In two dimensions the massive Schwinger model at zero Coleman angle is equivalent to the massive sine-Gordon theory[8] *and hence is equivalent to a neutral Yukawa gas.*[5]

For both Coulomb and Yukawa gases, singularities occur when opposite charges approach each other. In addition, there are self-interaction infinities. The self-energy terms can be eliminated by normal-ordering the potential, which is equivalent to absorbing the infinity into λ_0 as will be shown in the next section. This is well known to sine-Gordon theorists. The singularity resulting from plus-minus short-distance interaction is not so simply eliminated. One convenient possibility is to smear the point charges. This is a reasonable procedure since point charges never exist anyway. Replace a point charge at R_i by a charge distribution $f(r-R_i)$. Hence $\int f(x)d^3x=1$ and $f(x)$ is peaked about $x=0$. ρ in Eq. (2.3) would be replaced by

$$\rho_{\text{smeared}}(\vec{r})=\sum_{l=1}^{n}q_l f(\vec{r}-\vec{R}_l)\,. \qquad (2.10)$$

The limit $f(x)\to\delta^3(x)$ reproduces the point charge distribution. The effect on the field-theory representation is to replace $\exp[i\beta q\chi(R)]$ by

$$\exp\left[i\int \beta q\chi(\vec{r})f(\vec{r}-\vec{R})d^3r\right] \equiv \exp[i\beta(\chi * f)(\vec{R})]$$

so that $\cos\beta q\chi(\vec{R})$ of Eq. (2.4) gets replaced by $\cos\beta q(\chi * f)(\vec{R})$. This type of smearing is necessary in mathematical field theory where fields are distributions and must always be smeared with test functions. In these models the smearing of fields is natural since it corresponds to the smearing of point charges. The self-energy term also becomes finite and is equal to

$$\frac{q^2}{2}\int f(\vec{r})\,\frac{e^{-m|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|}\,f(\vec{r}')\,d^3r\,d^3r' \qquad (2.11)$$

for the Yukawa case. The self-energy in the Coulomb case is given by Eq. (2.11) with $m=0$ and in the general case is

$$\frac{q^2}{2}\int f(\vec{r})\,V(\vec{r},\vec{r}')f(\vec{r}')\,d^3r\,d^3r' \;. \qquad (2.12)$$

III. PERTURBATIVE VERIFICATION

The purpose of this section is twofold. First, the formal Gaussian representation is verified in perturbation theory. It is checked to third order in λ_0^3 for the Coulomb plasma model (sine-Gordon field theory) in three dimensions. All orders in β are resummed to give the first few terms of the grand partition function. *Thus perturbation theory when rearranged does indeed give the grand canonical sum.* The second purpose of this section is to set up a correspondence between perturbative Feynman diagrams and the statistical-mechanical system. This is done in the latter part of this section and the results are summarized in Table I.

$\langle \exp(2\lambda_0 \int \cos\beta q\chi)\rangle$ is the sum of vacuum bubble diagrams. To obtain the Feynman rules one could rescale χ so that $S_0 = \frac{1}{2}\int \nabla\chi\cdot\nabla\chi\, d^3R$; however, when not rescaled $S_0\ [=\int(\nabla\chi\cdot\nabla\chi/8\pi)d^3R]$ acts like the electrostatic energy of the system. To retain this physical meaning the Feynman rules will be listed without χ rescaled. For bubble diagrams they are as follows:

(a) Draw all topologically distinct vacuum bubbles (connected or disconnected) with vertices of an arbitrary even order (including zero order). Order, here, refers to the number of lines attached to a vertex.

(b) For each vertex associate a factor $(2\lambda_0)\int_V d^3r_i$. i refers to the ith vertex.

(c) For each vertex of order $2n$ associate a factor of $(-\beta^2 q^2)^n$.

(d) For each propagator associate a factor of $(1/\beta)(1/|r_i - r_j|)$.[9]

FIG. 1. Some Feynman graph combinatorial factors. In (a) there are four lines connecting the two vertices. According to rule (e) there is a factor of $1/4!$. (b) is a threefold self-energy tadpole. According to rule (f) there is a factor of $1/3!\,2^3$.

(e) Put in a factor $1/l!$ for each pair of vertices connected by l lines. See Fig. 1(a).

(f) Put in a factor of $1/(2l)!! = 1/2^l l!$ for each l-fold self-energy tadpole. See Fig. 1(b).

(g) Put in a factor of $[(\text{order of symmetry group of graph})!]^{-1}$.

(h) The empty graph is to be included and contributes unity.

Equivalent rules for (a) and (g) are as follows:

(a) Draw all bubble graphs (topologically distinct or not), that is, label the vertices and treat them as distinguishable.

(g) Put in a factor of $(\text{number of vertices!})^{-1}$.

The effect of the self-energy tadpoles is to renormalize λ_0. Any graph can be drawn as a graph without tadpoles plus tadpoles adjoined. Consider the effect of adding an arbitrary number of tadpoles to a "bare" vertex (see Fig. 2). The following factor will multiply the "tadpoleless" vertex:

$$\sum_{n=0}^{\infty}(-\beta^2q^2)^n\left(\frac{1}{n!}\,\frac{1}{2^n}\right)\left(\frac{1}{\beta|\vec{0}|}\right)^n = \exp\left(-\tfrac{1}{2}\beta q^2\,\frac{1}{|\vec{0}|}\right). \qquad (3.1)$$

If the smeared interaction $2\lambda_0\cos(\beta q\chi * f)$ is used Eq. (3.1) becomes

$$\exp\left[-\tfrac{1}{2}\beta q^2\int f(r)\,\frac{1}{|r-r'|}\,f(r')\,d^3r\,d^3r'\right]. \qquad (3.2)$$

One sees that the effect of tadpoles is to multiply each vertex by $e^{-\beta\times\text{self-energy}}$. Hence rules (b) and (f) are modified to the following:

(b) For each vertex associate a factor $2\lambda_R\int_V d^3r_i$ with λ_R the renormalized activity and $\lambda_R = \lambda_0 e^{-\beta\times\text{self-energy}}$.

(f) Do not include self-energy tadpole diagrams.

Alternatively one may use $2\lambda_R : \cos\beta q\chi :$ as the interaction density. Normal-ordering corresponds

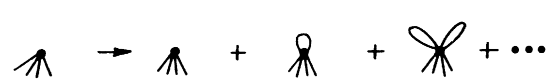

FIG. 2. The effect of self-energy tadpoles. The bare vertex is replaced by a sum of terms, each one with an additional tadpole attached.

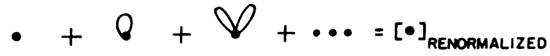

FIG. 3. The diagrams of order λ_0.

to a renormalization of λ_0. This fact, well known to sine-Gordon theorists, actually holds for any interaction which can be represented in Gaussian form. One can also see this in the grand canonical sum [Eq. (2.1)] where the self-energy terms would simply factor out to multiply $\lambda_0{}^n$ by $e^{-\beta n \times \text{self-energy}}$, i.e., $\lambda_0 \to \lambda_0 e^{-\beta \times \text{self-energy}} = \lambda_R$.

It will now be checked that to order $\lambda_R{}^3$ perturbation theory reproduces the $\mathfrak{z}$ of Eq. (2.1). To zeroth order the empty diagram contributes 1 and $\mathfrak{z}$ begins with 1. The diagrams of order λ_0 are shown in Fig. 3. They contribute $2\lambda_R \int_V d^3r = 2\lambda_R V$ which equals $\sum_{q_1=\pm q} \lambda_R \int_V d^3r$. The diagrams of order $\lambda_R{}^2$ are shown in Fig. 4. They sum to

$$\lambda_R{}^2 Z_2 = \tfrac{1}{2}(2\lambda_R)^2 \sum_{n=0}^{\infty} \frac{1}{(2n)!}(-\beta^2 q^2)^{2n} \times \int_V d^3R_1 d^3R_2 \frac{1}{(\beta|R_1-R_2|)^{2n}}$$
$$= \lambda_R{}^2 \int_V d^3R_1 d^3R_2 \left(e^{-\beta q^2/|R_1-R_2|} + e^{\beta q^2/|R_1-R_2|}\right) . \tag{3.3}$$

The $\frac{1}{2}$ factor multiplying the expression comes from the symmetry factor of rule (g). Equation (3.3) equals

$$\frac{\lambda_R{}^2}{2!} \sum_{q_1=\pm q} \sum_{q_2=\pm q} \int d^3R_1 d^3R_2 \exp\left(-\beta \frac{q_1 q_2}{|R_1-R_2|}\right) , \tag{3.4}$$

which is the second-order term of Eq. (2.1).

The diagrams of order $\lambda_R{}^3$ are shown in Fig. 5. They separate into two classes: those with an even number of propagators between vertices [Fig. 5(a)], and those with an odd number of propagators between vertices [Fig. 5(b)]. It is not too hard to sum these two sets to reproduce the third-order term in Eq. (2.1).

It is doubtless that perturbation theory reproduces the grand partition function to all orders in λ_R for this particular example, the Coulomb plasma in three dimensions. If another interaction had been used, or if several charges of different activities had been used, perturbation theory would have reproduced $\mathfrak{z}$. Of course, the Feynman rules would have to be modified. In particular, rule (d) would have to be replaced by the following:

(d) For each propagator associate a factor of $(1/\beta)\, V(r_i, r_j)$.

Particles of different activities would lead to vertices which would have to be distinguished. The appropriate λ's would have to be associated with the appropriate vertices, etc. Finally, the types of graphs and the factors associated with a particular vertex order would be different. The Feynman rules for other theories are thus easily obtained by modifying the rules presented in this paper.

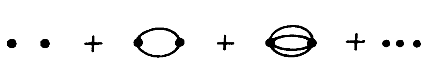

FIG. 4. The diagrams of order $\lambda_R{}^2$. The self-energy tadpoles are dropped according to the modified (b) and (f) rules.

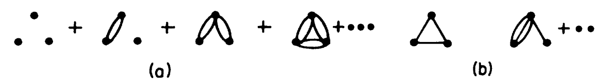

FIG. 5. Diagrams of order $\lambda_R{}^3$. Diagrams of (a) have an even number of propagators between vertices, whereas (b) have only an odd number.

A simplification can be made. There are zero-order vertices (vertices to which no lines are attached) in the Feynman graphs because the interaction $2\lambda_R : \cos\beta q\chi :$ when Taylor expanded begins with $2\lambda_R$. One can rewrite $2\lambda_R : \cos\beta q\chi :$ (or $\lambda^{(i)} e^{i\beta q^{(i)}\chi}$) as $2\lambda_R : (\cos\beta q\chi - 1) : + 2\lambda_R$ (similarly for $\lambda^{(i)} e^{i\beta q^{(i)}\chi}$ type terms). As a result the statement, "Do not include zero-order graphs but multiply all graphs by $e^{2\lambda_R V}$" is added to Feynman rule (a).

We know from perturbation theory that the contributions of graphs may be expressed in the form

$$\mathfrak{z} = \exp\left(\sum \text{connected graphs}\right)$$
$$= \exp\left[2\lambda_R V + \left(\sum_{l=2}^{\infty} \lambda_R{}^l b_l\right) V\right], \tag{3.5}$$

where b_l = the connected graphs of order $\lambda_R{}^l$. The volume factor is put into the definition of the b_l because vacuum bubble diagrams are proportional to V due to translational invariance (actually this is not quite correct because of boundary effects, but is approximately true in the large-volume limit). The vacuum energy per unit area (three dimensions = two space + one time) $\mathcal{E}$ for the field theory is

$$\mathcal{E} = \sum_{l=1}^{\infty} b_l \lambda_R{}^l \quad (b_1 \equiv 2) . \tag{3.6}$$

The b_l have an important connection with statistical mechanics. They are the cluster integrals of the Mayer expansion.[10] The thermodynamic properties are determined by these b_l. In particular,

$$\langle\rho\rangle=\sum_{l=1}^{\infty} l b_l \lambda_R{}^l\ , \qquad (3.7a)$$

$$\beta\langle p\rangle=\sum_{l=1}^{\infty} b_l \lambda_R{}^l\ , \qquad (3.7b)$$

where ρ is the density of particles, $\langle\rho\rangle=\langle N\rangle/V$, and p is the pressure. When λ_R is expressed in terms of $\langle\rho\rangle$ via Eq. (3.7a) and substituted into (3.7b) the equation of state is obtained. $\beta\langle p\rangle$ of the statistical-mechanical system is equal to the vacuum energy density $\mathcal{E}$ of the field theory.

The small parameter in the Mayer expansion is the function, $e^{-\beta V(r)}-1$. Often $V(r)$ is short-ranged so that $e^{-\beta V(r)}-1$ is nonzero only in a small region compared to V. Such a case occurs in the very massive ($m \gg 1/V$) Yukawa gas. An immediate application would be to the massive Schwinger model in two dimensions at zero Coleman angle, since, as previously noted, this is equivalent to the massive sine-Gordon Lagrangian. The charge e of the Schwinger model is related to the mass m of the massive sine-Gordon model by $m^2\pi/2=e^2$. Hence one may obtain results in the strong-coupling limit of the massive Schwinger model by using the Mayer cluster expansion applied to a Yukawa gas.[11]

In performing this check to order $\lambda_R{}^3$ one notices a correspondence between diagrammatic perturbation theory and the Coulomb gas. The vertices of Feynman diagrams are the ions and electrons of the plasma. In the Yukawa gas case, they would be the quanta and in the potential $[V(r)]$ case, one might call them the molecules. Up to a temperature factor the propagators represent the interactions. The number of interactions a particle undergoes is the same as the order of the vertex. Pairs of particles may undergo arbitrarily many interactions and when summed these give the Boltzmann factors. An external vertex at x corresponds to fixing a molecule in the gas at x. In field-theoretic language it is the vacuum expectation value of the operator $\Phi_q(x)\equiv e^{i\beta q\chi(x)}$. This operator may be interpreted as producing a charge q at x. Diagrams with several fixed external vertices are related to the correlation functions used in statistical mechanics.[7,10] In field theory they are the Green's functions of the $\Phi_q(x)$ fields. Finally writing

$$\mathfrak{z}=\sum_{N=0}^{\infty}\lambda_R{}^N Z_N$$

(with Z_N the partition function for N interacting particles), one sees that the Nth-order diagrams yield the N-particle partition function. Thus there is a complete correspondence between Feynman diagrams and the Coulomb plasma. This correspondence is summarized in Table I.

There is a sense in which the usual polynomial field theories (such as $g\chi^4/4!$) are Coulomb or Yukawa plasmas. In a high-temperature limit consider the sine-Gordon Lagrangian whose underlying statistical-mechanical system is the Coulomb plasma. The potential $V=-2\lambda_R:\cos(4\pi\beta)^{1/2}q\chi:$ [χ has been rescaled to eliminate the β dependence in H_0 ($=\frac{1}{2}\int\nabla\chi\cdot\nabla\chi$)] in this limit may be expanded in a Taylor series, $V=2\lambda_R[-1+4\pi\beta q^2:\chi^2/2:-(4\pi\beta q^2)^2:\chi^4/4!:+\cdots]$. If the system is "hot" and χ does not fluctuate violently from 0, then $V\approx-2\lambda_R+\lambda_R 8\pi\beta q^2:\chi^2/2:-2\lambda_R(4\pi\beta q^2)^2:\chi^4/4!:$ and hence one has a massive χ^4 theory with mass equal to $8\pi\lambda_R\beta q^2$ and a small negative χ^4 coupling constant g equal to $-32\pi^2\beta^2q^4\lambda_R$. Perturbation theory would be useful in this high-temperature limit since the coupling constant is small. Allowing arbitrary charges and activities the potential V becomes

$$V=-\sum_i \lambda^{(i)}e^{i\beta q^{(i)}\chi}\ . \qquad (3.8)$$

By adjusting the $\lambda^{(i)}$ and $q^{(i)}$ one may obtain better approximations to polynomial self-coupled field theories. In fact Eq. (3.8) is almost a Fourier transform. Unfortunately the $\lambda^{(i)}$ must be positive to retain their physical meaning. This restriction ruins the possibility of exact approximation.

TABLE I. Feynman graph correspondence.

Statistical mechanics		Field theory
particles	⟷	vertices
interactions	⟷	propagators
λ_0,β	⟷	coupling constants
Z_N	⟷	Nth-order diagrams
cluster expansion	⟷	expansion in $2\lambda_R\int_V(\cos\beta q\chi-1)$
correlation functions	⟷	Green's functions
a charge, q, produced at x	⟷	the operator $\Phi_q(x)=e^{i\beta q\chi(x)}$

IV. THE TWO-DIMENSIONAL SINE-GORDON THEORY

The last two sections have presented general methods and techniques. A specific example, the two-dimensional sine-Gordon theory,[12] will be, for the most part, the subject of the rest of this paper. As previously noted, this field theory is equivalent to a two-dimensional Coulomb gas. The interparticle potential is a logarithmic one:

$$V(r) = -2q_1q_2 \ln \frac{|R_1 - R_2|}{a_0}, \qquad (4.1)$$

with a_0 arbitrary. Equation (4.1) is also the interaction between two parallel lines of charge, one with a charge per unit length of q_1 and one of charge per unit length of q_2. Thus one may view the charges q as the charge densities in wires in the usual three-dimensional world. The wires are restricted to be perpendicular to a two-dimensional sheet. Another equivalent model is to replace the charged wires by currents. The magnetic interaction leads to the same logarithmic potential, $V = -2I_1I_2 \ln r/a_0$. The q's would then be the currents, I.

The two-dimensional Coulomb plasma differs from the three-dimensional version in one important way. Owing to infrared divergences, a smeared charge distribution has infinite energy unless it is neutral. Consider such a charge distribution ρ restricted to a finite region. The electric field goes like $Q_T{}^2 \vec{r}/r^2$ for large r. Therefore the energy density $\nabla\Phi \cdot \nabla\Phi/8\pi$ goes like $(Q_T{}^2/8\pi)(1/r^2)$ and hence the total energy diverges logarithmically unless the total charge Q_T is zero. In dealing with this two-dimensional Coulomb gas one has two choices. The first is to require total neutrality. $\mathfrak{z}$ would become

$$\mathfrak{z} = \sum_{n=0}^{\infty} \frac{\lambda_0{}^{2n}}{n!n!} \int d^2R_1 \cdots d^2R_n \int d^2X_1 \cdots d^2X_n \times \exp\left[\beta q^2 \sum_{i,j} \left(\ln \frac{|R_i - R_j|}{a_0} + \ln \frac{|X_i - X_j|}{a_0} - 2 \ln \frac{|R_i - X_j|}{a_0}\right)\right]. \qquad (4.2)$$

$\mathfrak{z}$ is independent of a_0 as long as self-energies are retained. $\mathfrak{z}$ still corresponds to the sine-Gordon field theory because non-neutral plasmas do not contribute to the functional integral; however, naive perturbation theory is incorrect. A correct way of obtaining the Feynman rules is to use the massive sine-Gordon Lagrangian and take the limit $m \to 0$. As soon as $m^2V \ll 1$ the massive propagator becomes $(2/\beta)\ln|mr|$. Perturbation theory when rearranged and partially summed gives Eq. (4.2) with $m = 1/a_0$. The non-neutral sums are proportional to $m^{\beta Q_T{}^2}$ and hence vanish as $m \to 0$. The Feynman rules should use the propagator of Eq. (4.2) in the limit where a_0 goes to infinity. The technique of adding a mass term to the sine-Gordon Lagrangian and letting the mass go to zero is not new. Coleman[13] used it in a paper showing the equivalence of the sine-Gordon theory with the massive Thirring model. It has a physical meaning since it demands neutrality of the plasma system. This paper will use this version of handling the infrared divergence. Total neutrality will always be maintained.

An alternative approach is to enclose the system in a grounded conducting casing. If there is an excess charge within V then an equal and opposite charge will appear on the conductor. In calculating the partition function one integrates $\nabla\Phi \cdot \nabla\Phi$ only over the volume V since the conductor causes $\nabla\Phi$ to be zero outside V. The Gaussian representation of $\mathfrak{z}$ would be modified to

$$\mathfrak{z} = \frac{1}{N} \iint \mathfrak{D}\chi \exp\left(-\beta \int_V \frac{\nabla\chi \cdot \nabla\chi}{8\pi} + 2\lambda_0 \int_V \cos\beta q \chi\right). \qquad (4.3)$$

The equation of motion, $(\beta/4\pi)\nabla^2\chi - 2\lambda_0 \beta q \sin\beta q \chi = 0$, is not valid because of surface terms due to integration by parts. These surface terms represent dynamical degrees of freedom and must be quantized. In principle this can be done using the techniques of Halpern and Senjanović.[14]

V. THE PHASES OF THE SINE-GORDON THEORY

This section will review the work done on the two-dimensional Coulomb gas[5,15] and relate it to the work done on the sine-Gordon field theory.[3,4,5,13,16,17] In particular, the phase of the system will be determined. Coleman has shown that a vacuum instability occurs when βq^2 gets too large.[13] This corresponds to a phase transition in the Coulomb system.[5]

Because the works of others will be referred to and because people have used different variables to denote the parameters of the sine-Gordon equation, there is some notational confusion. For example, the β that Coleman uses is not the inverse temperature. When confusion is possible I will subscript letters with the authors initials. For example, the β of Hauge and Hemmer is $\frac{1}{2}$ the β used in this paper so that β_{HH} will refer to their inverse temperature. $\beta_C = \beta_M = (4\pi\beta)^{1/2}q$ is the coupling constant used by Coleman[13] and Mandelstam.[16] This paper, for the most part, conforms with the notation of Kosterlitz and Thouless.

The method of Kosterlitz and Thouless will be used to determine the phases of the Coulomb system when λ_R is small. For $\beta q^2 \ll 1$, λ_R corresponds to the density, so that small λ_R means a dilute system. In fact, at $\beta q^2 = 0$,

$$\mathfrak{z} = \sum_{n=0}^{\infty} \frac{(\lambda_R V)^{2n}}{n!n!} .$$

Consider the situation where $\lambda_R V \gg 1$ even though $\lambda_R \ll 1$. $\mathfrak{z}$ has a maximum contribution for $n \gg 1$. Replacing the sum by an integral and using Stirling's formula,

$$\begin{aligned}\mathfrak{z} &\approx \int e^{2n\ln\lambda_R V + 2n - 2n\ln n - \ln n}\frac{dn}{2\pi} \\ &\approx \left[\frac{-1}{2\pi f''(n_{\max})}\right]^{1/2} e^{f(n_{\max})} \\ &\approx \left(\frac{\lambda_R V}{4\pi}\right)^{1/2} \exp[2\lambda_R V - \ln(\lambda_R V)] \\ &= \left(\frac{1}{4\pi\lambda_R V}\right)^{1/2} e^{2\lambda_R V} ,\end{aligned} \tag{5.1}$$

where $f(n) = 2n\ln\lambda_R V + 2n - 2n\ln n - \ln n$ and $n_{\max} e^{1/2n_{\max}} = \lambda_R V$. $n_{\max} \approx \lambda_R V$ for $\lambda_R V$ large. The integral has been approximated by Laplace's method,

$$\langle N\rangle = \lambda_R \frac{\partial}{\partial\lambda_R}\ln\mathfrak{z} \approx 2\lambda_R V, \quad \lambda_R = \frac{\langle N\rangle}{2V}$$

which shows that λ_R is indeed a density. The limits $\lambda_R V \gg 1$ but $\lambda_R \ll 1$ correspond to a situation where many ions are present but the density is small which is the proper statistical limit (in the limit $\lambda_R V \ll 1$, $\langle N\rangle \approx \frac{1}{2}\lambda_R^2 V^2$ so $\langle N\rangle \ll 1$ which is undesirable).

When λ_R is small one can calculate the mean square distance between an ion and an elctron by assuming that the other charges in the plasma may be neglected. In fact the exact expression for the mean square distance is

$$\langle r^2\rangle = \frac{\displaystyle\sum_{N=0}^{\infty} \frac{\lambda_0^{2N}}{N!N!}\int d^2X_1\cdots d^2X_N \int d^2R_1\cdots d^2R_N (X_1 - R_1)^2 e^{-\beta U_N}}{\displaystyle\sum_{N=1}^{\infty} \frac{\lambda_0^{2N}}{N!N!}\int d^2X_1\cdots d^2X_N \int d^2R_1\cdots d^2R_N e^{-\beta U_N}}, \tag{5.2}$$

where U_N is the energy of the configuration [see Eq. (4.2)]. The $N=1$ term gives for the mean square distance

$$\langle r^2\rangle = \frac{(R^{4-2\beta q^2} - r_0^{4-2\beta q^2})(2 - 2\beta q^2)}{(4 - 2\beta q^2)(R^{2-2\beta q^2} - r_0^{2-2\beta q^2})} , \tag{5.3}$$

where r_0 is an ultraviolet cutoff introduced to make $\langle r^2\rangle$ well defined for $\beta q^2 \geq 1$. In fact, if the charges were not point charges but "ringlets" of charge densities of radius r_0 Eq. (5.3) would be the mean square distance between ringlets. In using Eq. (5.3) I am not implying that the $N=1$ term dominates. It is obvious from the above discussion that a large value of N dominates. Using the dominate term is, of course, like using a partition function in lieu of the grand sum. Equation (5.3) is inaccurate in the region $r > (V/\langle N\rangle)^{1/2}$ and the integrals should probably be cut off at such a value (which is still a large number in the dilute-gas approximation). In Eq. (5.3) I have neglected "edge effects" which occur if one of the charges is near the boundary of the volume. Equation (5.3) is calculated on the basis that one of the charges is at the center of the volume.

Equation (5.3) yields the following result for $\beta q^2 < 1$:

$$\langle r^2\rangle \approx \frac{R^2(1 - \beta q^2)}{(2 - \beta q^2)} \;(\lambda_R \text{ small}) . \tag{5.4}$$

The fluctuation in the distance between charges according to Eq. (5.4) is large and hence for $\beta q^2 < 1$ the Coulomb system is in the plasma phase, that is, the electrons and ions do not pair up to form dipoles. At $\beta q^2 = 1$ the same is still true since

$$\langle r^2\rangle \approx \frac{R^2}{2\ln(R/r_0)} \;(\lambda_R \text{ small}) . \tag{5.5}$$

In the region where $1 < \beta q^2 < 2$, $\langle r^2\rangle$ is renormalization dependent, i.e., $\langle r^2\rangle$ depends on the parameter r_0:

$$\langle r^2\rangle \approx \frac{\beta q^2-1}{2-\beta q^2} R^2 \left(\frac{r_0}{R}\right)^{2\beta q^2-2}$$
$$= \frac{\beta q^2-1}{2-\beta q^2} r_0{}^2 \left(\frac{R}{r_0}\right)^{4-2\beta q^2} . \tag{5.6}$$

Equation (5.6) shows that although $\langle r^2\rangle$ is not of the order of the size of the system, it is still much greater than $r_0{}^2$ and hence a dipole collapse has not yet occurred. Although some dipoles may exist, the preponderance of ions and electrons is still unbound and thus the Coulomb plasma phase still occurs. Of course as λ_R gets larger the above conclusions may no longer be correct since Eq. (5.3) is based on the dilute-gas approximation. It is conceivable that if λ_R gets large enough a phase transition might occur.

Finally for $\beta q^2=2$ and $\beta q^2>2$ the dipole phase occurs since

$$\langle r^2\rangle \approx 2 r_0{}^2 \ln\frac{R}{r_0} \quad (\beta q^2=2) ,$$
$$\langle r^2\rangle \approx \frac{\beta q^2-1}{\beta q^2-2} r_0{}^2 \quad (\beta q^2>2) . \tag{5.7}$$

A phase transition occurs around $\beta q^2=2$ for small λ_R. The nature of the phase transition is simple: As βq^2 increases the free ions and electrons of the Coulomb plasma collapse to form dipoles, and a new gas of weakly interacting dipoles is formed. This phase transition has been examined in more detail by Kosterlitz and Thouless, who find a divergence in the polarizability in going from high β to low β (i.e., from dipoles to the plasma). This divergence is understandable since as the temperature increases the average separation between a plus charge and minus charge forming a dipole increases. This causes the dipole moment to increase and as a result the polarizability of the system is greatly enhanced. Using reasonable methods they obtain that $q^2\beta_{\text{critical}} \approx 2$.

This phase transition has an important relation to the sine-Gordon field theory. The point $\beta q^2=2$ corresponds to $\beta_C{}^2=8\pi$. It was precisely at this point that Coleman found a vacuum instability. *One can now understand this instability from the Coulomb point of view: It is precisely a phase transition from an ion plasma to a dipole gas.*

Solitons are known to exist as solutions to the sine-Gordon equation.[18] The way the critical temperature varies as λ_R varies is important since it may affect the number and stability of soliton-antisoliton bound states. Luther[17] has proved quantum mechanically that stable bound states occur for $n=1, 2, \ldots, \beta q^2/(2-\beta q^2)$ $(0<\beta q^2<2)$ ($n=0$ always exists and is the usual soliton) with masses of the form

$$m_n = C(\beta) \sin\frac{n\pi}{2} \frac{\beta q^2}{2-\beta q^2} , \tag{5.8}$$

where $C(\beta)$ is a temperature-dependent renormalized constant. The renormalization of C depends on the lattice spacing and the x-y anisotropy (Luther used the spin-$\frac{1}{2}$ x-y-z lattice chain to obtain the above results). This seems to indicate that the number of bound states does not depend on λ_R.

There are three possible phase diagrams which might occur. These are shown in Fig. 6. In these diagrams the pressure p or the density ρ may be substituted for λ_R if the equation of state is known. One strongly suspects that λ_R increases monotonically as p or ρ increases for fixed volume and temperature. If Fig. 6(a) is the situation, the dipole phase (λ_R very large) will probably prevent the solitons from existing even though βq^2 is less than 2. Both (a) and (c) cases require that an additional parameter enter the theory, since λ_R, being the only dimensional constant, has no dimensional quantity to set the scale.[19] It is therefore meaningless to plot λ_R versus βq^2. Indications (see Sec. VII) are that for $\beta q^2 \geq 1$ a cutoff must be introduced due to ultraviolet singularities. If this cutoff can be removed without introducing a new dimension into the theory (i.e., there is no dimenional transmutation) then the only possibility would be (b) and solitons of arbitrary n occur for βq^2 sufficiently close to 2. Kosterlitz and Thouless obtained a curve similar to (c). They got $\beta_{\text{critical}}\, q^2 = 2(1+c\lambda_R)$ with $c \approx 1.3\pi$. They used a cutoff in their potential $[U_{KT}(r) = -q_1 q_2 \ln(r/r_0) + 2\mu$ for $r>r_0$ and $U(r)=0$ for $r>r_0]$. The reason for their result is simple: They view the Coulomb system from the dipole side of $\beta_{\text{critical}}\, q^2$. Their first approximation was to neglect the effects of all other dipoles in calculating $\langle r^2\rangle$, the mean distance squared between the plus and minus constituents of the dipole. They obtained $\beta_{\text{critical}}\, q^2=2$. What corrections result if other dipoles are taken into consideration? Basically, it will be easier to separate the plus and minus constituents because dipoles will interpolate to reduce the potential. Consequently, each charge is partially screened and it will be easier to pull them apart. The presence of dipoles lowers

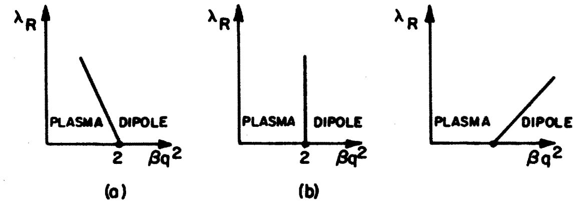

FIG. 6. Three possible phase diagrams.

the temperature at which the phase occurs (in other words $\beta_{\text{critical}}\, q^2$ increases with λ_R). They expressed the screening in terms of an effective dielectric constant $\epsilon(r)$ which depended on the separation of the plus and minus. If possibility (c) where to occur as Kosterlitz and Thouless have predicted, one would expect the number of states to be different from the current prediction (a natural guess would be $n = 1, 2, \ldots, \beta/(\beta_{\text{critical}} - \beta)$ and correspondingly $m_n = C(\beta)\sin(n\pi/2)\,[\beta/(\beta_{\text{critical}} - \beta)]$). I refer the reader to their paper for their results. It may be that the situation depends on how one modifies the potential at short distances (to eliminate the ultraviolet singularities), and thus Kosterlitz and Thouless's result is one possible example which might occur.

VI. THE NONLINEAR σ MODEL

This section will show the equivalence of a nonlinear σ model and the sine-Gordon equation.[20] First consider the linear O(2) σ model with a linear symmetry-breaking term,

$$\mathcal{L} = \tfrac{1}{2}(\partial\sigma)^2 + \tfrac{1}{2}(\partial\Pi)^2 - V(\sigma, \Pi) , \tag{6.1}$$

and

$$V(\sigma, \Pi) = -2a\sigma + g(\sigma^2 + \Pi^2 - f_\Pi^{\,2}) . \tag{6.2}$$

When $a = 0$, $m_\Pi^{\,2} = 0$ (Goldstone boson) and $m_\sigma^{\,2} = 16 g f_\Pi^{\,2}$ (the usual partially conserved axial-vector current type relation of the σ mass being proportional to f_Π). When $a \neq 0$ the minimum of V occurs at $\Pi = 0$,

$$\sigma = (\text{sign } a) f_\Pi + \frac{a}{4 f_\Pi^{\,2}}\left(\frac{1}{g}\right) + O\left(\frac{1}{g^2}\right)$$

and the Π field acquires a mass,

$$m_\Pi^{\,2} = \frac{2|a|}{f_\Pi} + O\left(\frac{1}{g}\right) .$$

The nonlinear σ model with a linear symmetry-breaking term is the limit of Eq. (6.2) as $g \to \infty$. This has the effect of requiring $\sigma^2 + \Pi^2 = f_\Pi^{\,2}$. In fact, if Eq. (6.1) were used in a functional integral the limit $g \to \infty$ would produce a functional δ function, $\delta(\sigma^2 + \Pi^2 - f_\Pi^{\,2})$. To enforce the relation $\sigma^2 + \Pi^2 = f_\Pi^{\,2}$ let $\sigma = -f_\Pi \cos\theta$ and $\Pi = -f_\Pi \sin\theta$ where θ is a new field. $\mathcal{L}$ becomes

$$\mathcal{L} = \tfrac{1}{2} f_\Pi^{\,2}(\partial\theta)^2 + 2 a f_\Pi \cos\theta , \tag{6.3}$$

which is the sine-Gordon Lagrangian. Rescaling $\theta = (1/f_\Pi)\chi$ gives

$$\mathcal{L} = \tfrac{1}{2}(\partial\chi)^2 + 2 a f_\Pi \cos\left(\frac{1}{f_\Pi}\chi\right) , \tag{6.4}$$

or $V(\chi) = -2 a f_\Pi \cos[(1/f_\Pi)\chi]$ from which one can translate the parameters of this model with the Coulomb gas parameters:

$$\lambda_0 = 2 a f_\Pi ,$$
$$(4\pi\beta)^{1/2} q = \frac{1}{f_\Pi} , \quad \text{or } \beta q^2 = \frac{1}{4\pi f_\Pi^{\,2}} . \tag{6.5}$$

Since a phase transition takes place in the sine-Gordon field theory at $\beta q^2 \approx 2$ for λ_R small, *a phase transition must occur in the nonlinear O(2) σ model with a small symmetry-breaking term for $f_\Pi \approx 1/\sqrt{8\pi}$*. This complements the results of Brêzin and Zinn-Justin[21] and Bardeen, Lee, and Shrock,[22] who find transitions in the $O(N)$ nonlinear σ model in $2 + \epsilon$ dimensions for $N > 2$.

VII. RENORMALIZATION

Coleman[13] has indicated that the *only* renormalization necessary in the two-dimensional sine-Gordon model is for the self-energy tadpoles [such as in Fig. 1(b)]. These infinite contributions can be absorbed in λ_0. One should use a renormalized activity λ_R and ignore the tadpole self-energies. Coleman's result is true for $\beta q^2 < 1$ ($\beta_C^{\,2} < 4\pi$). For $\beta q^2 \geq 1$ the result is incorrect: Although there are no divergent graphs in any finite order (in λ_R and β) of perturbation theory, there *are* divergences when graphs are summed. The reader has already seen an example of this: The connected vacuum bubbles of order $\lambda_R^{\,2}$ (the graphs of Fig. 4 minus the first one). They sum to

$$2\lambda_R^{\,2} V \int_V d^2 r\,(r^{-2\beta q^2} - 1) ,$$

which converges for $\beta q^2 < 1$ and diverges for $\beta q^2 \geq 1$. Vacuum bubbles are not the only diagrams with divergences which cannot be absorbed in λ_0. Consider the contributions to the two-point function from the graphs of Fig. 7. They sum to

$$S(x,y) = (2\lambda_R)^2 \int_V d^2 z_1 \int_V d^2 z_2 \ln\frac{|x - z_1|}{a_0} \ln\frac{|z_2 - y|}{a_0}\left[\left(\frac{|z_1 - z_2|}{a_0}\right)^{2\beta q^2} - \left(\frac{|z_1 - z_2|}{a_0}\right)^{-2\beta q^2} - 1\right] . \tag{7.1}$$

The parameter a_0 has been left in and naive Feynman rules have been used in calculating Eq. (7.1). This is because the Green's functions for the fields $\chi(x)$ are ill defined. The interesting and well-defined operators are the $\Phi_q(x) = e^{i\beta q \chi(x)}$. Equation (7.1) converges for $\beta q^2 < 1$ and diverges for $\beta q^2 \geq 1$ [the same would be true if one calculated the Green's functions for the $\Phi_q(x)$]. Note that

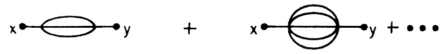

FIG. 7. Contributions to the two-point function. Although individual graphs are not divergent, the sum of these graphs gives a divergent contribution to the two-point function when $\beta q^2 \gtrsim 1$.

any individual graph of Fig. 4 or Fig. 7 is convergent. *Thus one has a situation* ($\beta q^2 \geq 1$) *where to any finite order (in* βq^2*) graphs have no ultraviolet divergences yet when perturbation theory is summed to all orders ultraviolet divergences appear, implying that nonperturbative renormalization methods are necessary.* Of course, one can look at $\mathfrak{Z}$ [Eq. (4.2)] directly to see that there are ultraviolet divergences for $\beta q^2 \geq 1$.

This section will consist of showing that to all orders in λ_R (as well as βq^2) $\mathfrak{Z}$ is well defined when $\beta q^2 < 1$. I take this to be a proof that *the O(2) nonlinear σ model and the sine-Gordon field theory are renormalizable to all orders for* $\beta q^2 < 1$.[23] For $\beta q^2 \gtrsim 1$ a cutoff must be introduced. Whether it may be removed by wave-function and coupling-constant renormalizations is unknown. If additional interaction counterterms must be added to the theory, then the sine-Gordon Lagrangian would have to be modified and the equation of motion would be an inaccurate representation. An implication of this: The soliton-antisoliton doublets would not exist. I suspect that renormalization should be possible at least for $1 < \beta q^2 < 2$ since lattice methods[17] have shown the existence of these doublets. It is still uncertain how to do this in the continuum field theory, although the equivalence of the nonlinear σ model and the sine-Gordon theory offers a possibility: Since g has dimensions of $(\text{mass})^2$ and the linear σ model [Eqs. (6.1) and (6.2)] is renormalizable, g acts as a cutoff for the sine-Gordon theory. If one could show that the relevant quantities are g independent (or g dependence can be absorbed into λ_0) for large g then this would provide the method of renormalization. The same is true in dimensions three and four where the equivalence between the two models still holds. Thus there is the possibility that using the linear σ model one can renormalize the sine-Gordon Lagrangian in three and four dimensions.

I will now present strong evidence that the partition function for the two-dimensional Coulomb gas [Eq. (4.2)] converges for $\beta q^2 < 1$. λ_R, of course, is $e^{-\beta \times \text{self-energy}}$ absorbed in λ_0. The method is not intended to be rigorous: Physical arguments are used to approximate $\mathfrak{Z}$. Equation (4.2) contains only neutral configurations because of the infrared singularity of the theory as discussed in Sec. IV. The arguments of this section also apply to the situation where neutrality is not required.

$\mathfrak{Z}$ acquires a large contribution whenever an x_i approaches a y_j and βq^2 is near 1. The nature of the singularity is governed by

$$VI(\beta q^2, \epsilon) = \int d^2x \int_{|x-y|<\epsilon} d^2y\, e^{-2\beta q^2 \ln|x-y|} \approx V \frac{\pi \epsilon^{2-2\beta q^2}}{1-\beta q^2}\,, \tag{7.2}$$

Equation (7.2) is the contribution to $\mathfrak{Z}$ when a plus charge and minus charge are within ϵ of each other (boundary effects being neglected). If ϵ is sufficiently small the plus-minus dipole will look like a neutral object and will interact very weakly with other charges and dipoles, even when these objects approach the dipole. ϵ is just a small parameter. If a plus charge and minus charge are within ϵ of each other then one says a dipole is present in the system. Since there will be a mean density ρ for the plasma and in this dissociated phase charges are randomly located, the average distance between charges is roughly $(1/\rho)^{1/2}$. One can take ϵ to be a fraction of this distance, say $\epsilon = 0.1(1/\rho)^{1/2}$.

Consider the term in $\mathfrak{Z}$ with N plus charges and N minus charges (i.e., Z_N) and single out the contribution due to dipoles:

$$Z_N = \frac{1}{N!N!}\Bigg[\underset{|x_i - y_j|<\epsilon \text{ for some } i,j}{\int d^2x_1 \cdots d^2x_N \int d^2y_1 \cdots d^2y_N\, e^{-\beta U_N}} + \underset{|x_{i_1}-y_{j_1}|<\epsilon,\ |x_{i_2}-y_{j_2}|<\epsilon \text{ for some } i_1,j_1,i_2,j_2}{\int d^2x_1 \cdots d^2x_N \int d^2y_1 \cdots d^2y_N\, e^{-\beta U_N}} + \cdots$$
$$+ \underset{\substack{|x_{i_1}-y_{j_1}|<\epsilon,\ |x_{i_2}-y_{j_2}|<\epsilon,\ \ldots,\ |x_{i_N}-y_{j_N}|<\epsilon \\ \text{for some ordering of the } x\text{'s and } y\text{'s}}}{\int d^2x_1 \cdots d^2x_N \int d^2y_1 \cdots d^2y_N\, e^{-\beta U_N}} + \underset{\text{no dipoles}}{\int d^2x_1 \cdots d^2x_N \int d^2y_1 \cdots d^2y_N\, e^{-\beta U_N}} \Bigg]. \tag{7.3}$$

The first N terms in Eq. (7.3) have precisely $1, 2, \ldots, N$ dipoles. The last term is the no-dipole term where no x_i is within ϵ of any y_j. Because dipoles interact weakly they may be factored out of the summands of Eq. (7.3):

$$Z_N \approx \sum_{l=0}^{N} \frac{V^l [I(\beta q^2, \epsilon)]^l Z_{N-l}^{\text{no dipole}}(\epsilon)}{l!}, \tag{7.4}$$

where

$$Z_n^{\text{no dipole}} = \frac{1}{n!}\frac{1}{n!}\int d^2x_1\cdots d^2x_n \int d^2y_1\cdots d^2y_n\, e^{-\beta U_n(x_i, y_j)}, \quad \text{no } x_i \text{ within } \epsilon \text{ of } y_j\,. \tag{7.5}$$

The combinatorial factor is accounted for as follows: There are

$$\frac{[N(N-1)\cdots(N-l+1)]^2}{l!}$$

ways of pairing up l pluses with l minuses from a set of N pluses and N minuses. This number times $(1/N!)(1/N!)$ gives the

$$\frac{1}{l!(N-l)!\,(N-l)!}$$

of Eqs. (7.4) and (7.5). Equations (4.2), (7.3), and (7.4) yield

$$\mathfrak{z} \approx e^{VI(\beta q^2,\, \epsilon)\lambda_R{}^2}\, \mathfrak{z}^{\text{no dipole}}(\epsilon)\,. \tag{7.6}$$

As long as $\beta q^2 < 1$, $I(\beta q^2, \epsilon)$ is small as one sees from Eq. (7.2). It remains to show $\mathfrak{z}^{\text{no dipole}}(\epsilon)$ does not diverge. Heuristically the reason for this is as follows: $\mathfrak{z}^{\text{no dipole}}$ vanishes if a plus approaches a plus or a minus approaches a minus and hence the charges must be evenly distributed when N becomes large. Consider a minus charge. Since the plasma is neutral it will see a charge distribution of +1. Neglecting boundary effects, one can lump this +1 charge distribution at some effective distance $r_{\text{eff}}(N, \epsilon)$. For ϵ small enough r_{eff} will be independent of ϵ and for N large enough it will be a slowly varying function of N. For simplicity take r_{eff} to be a constant as N goes to infinity. Then for large N

$$Z_N^{\text{no dipole}} \sim \frac{1}{N!}\frac{1}{N!}(r_{\text{eff}}{}^{-2\beta q^2})^{2N} V^{2N} \tag{7.7}$$

and this implies

$$\mathfrak{z}^{\text{no dipole}} \equiv \sum_{n=0}^{\infty} \lambda_R{}^{2N} Z_N^{\text{no dipole}}$$

converges.

To make the above argument more precise let N be large. Break V into $2N$ square cells of volume $V/2N$. The length of the side of a cell is $(V/2N)^{1/2} = d(N)$. Approximate $Z_N^{\text{no dipole}}$ by summing over all ways of placing the plus and minus charges into the $2N$ cells:

$$Z_N^{\text{no dipole}} \approx \sum_C \frac{1}{N!\,N!}\left\{\left[\left(\frac{V}{2N}\right)^{1/2}\right]^2\right\}^{2N} e^{-\beta U(C)}\,. \tag{7.8}$$

$[(V/2N)^{1/2}]^2$ is the area in which a charge is allowed to roam, C is a placement of the pluses and minuses into cells, and $U(C)$ is the energy of such a configuration (calculated with the charges at the center of the cells). The minimum energy configuration by symmetry occurs when the plus and minus charges alternate as in Fig. 8. To approximate the energy of this configuration, pick a charge. It has four nearest opposite charges [Fig. 9(a)] contributing a factor of $(d^4)^{-2\beta q^2}$ and four nearest like charges [Fig. 9(b)] contributing a factor of $[(\sqrt{2}\,d)^4]^{2\beta q^2}$. In the next row [Fig. 9(c)] there are eight opposite charges and eight like charges giving a factor of

$$[(2d)^4]^{2\beta q^2}[(\sqrt{5}\,d)^8]^{-2\beta q^2}[(\sqrt{8}\,d)^4]^{2\beta q^2}\,.$$

To this order

$$e^{-\beta U(\text{single charge})} = [4(\tfrac{4}{5})^2(\tfrac{8}{5})^2]^{2\beta q^2}\,.$$

Taking into consideration all rows and neglecting edge effects

$$\begin{aligned} f(N) &= e^{-\beta U(\text{single charge})} \\ &\approx \prod_{n=0}^{\sqrt{N}}\prod_{m=1}^{\sqrt{N}}[(n^2+m^2)^{1/2}]^{4(-1)^{m+n}2\beta q^2} \\ &= \exp\left[\sum_{n=0}^{\sqrt{N}}\sum_{m=1}^{\sqrt{N}}(-1)^{n+m}4\beta q^2\ln(n^2+m^2)\right]. \end{aligned} \tag{7.9}$$

Because the signs alternate f will be a slowly varying function of N (for example the third row multiplies the result of the first two by only $[(\frac{10}{9})^2(\frac{10}{14})^2(\frac{18}{14})^2]^{2\beta q^2} \approx (1.04)^{2\beta q^2}$). The total contribution to $e^{-\beta U(C_{\min})}$ is $[f(N)^{2N}]^{1/2}$ ($2N$ for each particle and a $\frac{1}{2}$ for double counting). One can

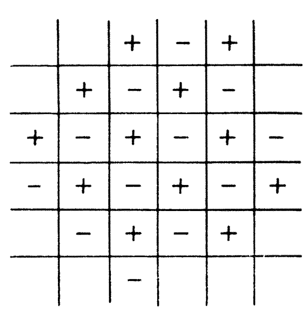

FIG. 8. Minimum-energy configuration.

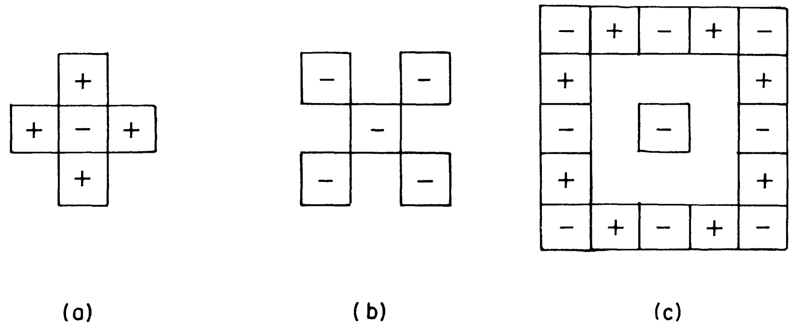

FIG. 9. Leading contributions to the energy of a minus charge.

bound $Z_N^{\text{no dipole}}$ by replacing $e^{-\beta U}$ by its maximum value $f(N)^N$. Since there are $(2N)!$ ways of putting the charges into the cells,

$$Z_N^{\text{no dipole}} < \frac{(2N)!}{N!\,N!}[f(N)]^N\left(\frac{V}{2N}\right)^{2N} \sim \frac{V^{2N}f(N)^N(4\pi N)^{1/2}e^{-2N}}{N!\,N!}, \qquad (7.10)$$

where Stirling's approximation has been used in the last step. Since $f(N)$ [Eq. (7.9)] is a slowly varying function of N, $Z_N^{\text{no dipole}}$ is highly attenuated for N large and $\mathfrak{Z}_{\text{no dipole}}$ converges.

It has now been shown that $\mathfrak{Z}(\lambda_R, \beta q^2)$ is finite for $\beta q^2 < 1$. $\mathfrak{Z}$ represents the sum of all vacuum bubble diagrams (in a finite-volume limit). Since this sum converges this indicates that other functions (such as Green's functions of the relevant operators) will have no divergences. It is folklore that the vacuum bubbles represent the most ultraviolet-divergent graphs.

The main result of this section is that the prescription of absorbing the self-energies into the activity renormalizes the sine-Gordon theory to all orders when $\beta q^2 < 1$.

For the statistical analog field theories discussed in this paper, renormalization may be regarded as the removal of smearing functions. Consider the sine-Gordon theory. When charges are smeared by an appropriate f, a well-defined nontrivial partition function $\mathfrak{Z}(f)$ is obtained (well defined in the sense that no infinities occur and nontrivial in the sense that $\mathfrak{Z}(f)$ is not 1 or $e^{2\lambda_R V}$, which is the ideal gas grand partition function). $\mathfrak{Z}(f)$ is, however, nonrelativistic. One would like to take the limit $f(x) \to \delta^d(x)$ so as to recover Poincaré invariance. Doing this naively causes $\mathfrak{Z}(f)$ to go to one because the $N \neq 0$ terms give zero due to the infinite self-energy. The way to avoid this problem is to let λ_0, the bare activity, depend on f: $\lambda_0 \equiv \lambda_0(f)$. $\mathfrak{Z}$ becomes

$$\mathfrak{Z}(f) = \frac{1}{N}\iint \mathfrak{D}\chi \exp\left[-\beta\int \frac{\nabla\chi\cdot\nabla\chi}{8\pi} + 2\lambda_0(f)\int_V \cos\beta q(\chi * f)\right]. \qquad (7.11)$$

Equation (7.11) gives an upper bound on $\mathfrak{Z}$ of $e^{2\lambda_0(f)V}$ since the functional $\int_V \cos\beta\chi * f \leq V$. $\mathfrak{Z}$ has a lower bound of 1 since the partition function is a sum of positive terms beginning with 1. $\lambda_0(f)$ is adjusted so that $\lim_{f\to\delta^d}\mathfrak{Z}(f) \equiv \mathfrak{Z}$ is nontrivial and any such limit should produce a good theory. The reader has seen one example of this, the sine-Gordon theory in two dimensions in the region $\beta q^2 < 1$. If one takes

$$\lambda_0(f) = C\exp\left[-\beta q^2\int f(r)\ln|r - r'|\,f(r')\,d^2r\,d^2r'\right], \qquad (7.12)$$

where C is arbitrary (from $\lambda_R = \lambda_0 e^{-\beta\times\text{self-energy}}$ one identifies C with λ_R), then the limit $f(x) \to \delta^2(x)$ produces a well-defined $\mathfrak{Z}$. $\lambda_0(f)$ goes to infinity in such a way as to keep λ_R finite and prevent $\mathfrak{Z}$ from going to 1. In the region $\beta q^2 \geq 1$, the $\lambda_0(f)$, defined by Eq. (7.12), would cause $\mathfrak{Z}(f)$ to become infinite as $f \to \delta^2$ and in this new region one must not let $\lambda_0(f)$ go to infinity as fast as in the $\beta q^2 < 1$ region. Possibly a sequence of $\lambda_0(f)$'s can be found which affects a cancellation between self-energy and interparticle interaction infinities.

Green's functions (and other relevant objects) must always be calculated using this limiting procedure. Consider

$$G(x,y) = \langle e^{i\beta q\chi(x)}e^{-i\beta q\chi(y)}\rangle = \frac{1}{N}\iint \mathfrak{D}\chi \exp\left(-\beta\int\frac{\nabla\chi\cdot\nabla\chi}{8\pi} + 2\lambda_0\int_V \cos\beta q\chi\right) \times e^{i\beta q\chi(x)}e^{-i\beta q\chi(y)}. \qquad (7.13)$$

Even after smearing ($\int \cos\beta q\chi * f$) and using $\lambda_0(f)$ of Eq. (7.12), this plus-minus Green's function is zero due to the ultraviolet infinite self-energy produced by $e^{i\beta q\chi(x)}$ and $e^{-i\beta q\chi(x)}$. The correct way of calculating is to replace Eq. (7.13) by

$$G(x,y)=\frac{1}{N}\iint \mathfrak{D}\chi \exp\left[-\beta\int\frac{\nabla\chi\cdot\nabla\chi}{8\pi}+2\lambda_0(f)\int_V\cos\beta q(\chi*f)\right]\times Z^2(f)\,e^{i\beta q(\chi*f)(x)}e^{-i\beta q(\chi*f)(y)}. \tag{7.14}$$

where $Z(f)$ is a wave-function renormalization constant. From physical principles one knows $Z(f)$ must be proportional to

$$\exp\left[-\beta q^2\int f(r)\ln|r-r'|f(r')d^2r\,d^2r'\right]$$

since $G(x,y)$ has a physical interpretation: $G(x,y)\mathfrak{z}$ is the partition function for a neutral Coulomb gas with a plus charge at x and a minus charge at y.

Nonrenormalizability can be viewed as follows: As $f\to\delta^d$ dipoles, triatomic molecules, molecular rings, and other polyatomic structures will begin to form. One must introduce renormalized activities for each of these structures. If $f\to\delta^d$ is too singular a limit to take, an infinite number of polyatomic objects will form causing one to introduce an infinite number of renormalized activities. This infinite set is reminiscent of what happens with polynomial field theories such as χ^6 in four dimensions. Here one is forced to introduce an infinite set of counterterms, $\sum_{n=1}^{\infty}(\delta g_{2n})\chi^{2n}$. The δg_{2n} couplings correspond to the unrenormalized activities. In the high-temperature limit of the sine-Gordon theory

$$V(\chi)\approx 2\lambda_R\left[-1+(4\pi\beta q^2)\frac{\chi^2}{2}-(4\pi\beta q^2)^2\frac{\chi^4}{4!}+(4\pi\beta q^2)^3\frac{\chi^6}{6!}\right],$$

which is a polynomial field theory with a χ^6 leading term. The high-temperature sine-Gordon theory being similar to this polynomial potential implies the formation of polyatomic structures in the χ^6 theory. Thus one suspects that the cause for nonrenormalizability for the statistical analog field theories is the same as in nonrenormalizable polynomial field theories.

VIII. TIDBITS

The correspondence between sine-Gordon theory and the Coulomb gas indicates that many effects are completely missed in the naive treatment of the theory. Most important is the structure of the vacuum. When λ_R is very small perturbation theory is valid. The vacuum looks like a "vacuum" since few charges are present. When λ_R gets larger, the vacuum is full of charges and the perturbation theory vacuum is a poor approximation to the real vacuum which contains many plus and minus ions. Also missed in perturbation theory is the phase transition at βq^2 near 2. When βq^2 is small the vacuum (and hence the entire theory) is radically different from when βq^2 is large. The importance of the nature of the vacuum is neglected in most treatments of field theory. It will be a complicated vacuum structure which will lead to quark confinement, asymptotic freedom, and the hadron spectrum. A vacuum consisting of a gas of "quanta" would be compatible with asymptotic freedom and quark confinement. When two quarks are placed at small separation distance usually no vacuum quanta will be between them. The physical vacuum has little effect on these two quarks. Hence small-distance behavior would be governed by bare vacuum and free interactions. At large separation distance the quanta would interpolate between the quarks and strange effects could occur. For example, a condensation into another phase might take place in the region between quarks especially if the gas is near a phase transition point. Such a condensation could provide a confining potential. In the two-dimensional sine-Gordon theory in the plasma phase (βq^2 small) one has asymptotic freedom. $G(x,y)$ [Eq. (7.14)] goes like $|x-y|^{-2\beta q^2}$ for small $|x-y|$. This is the same as in a free field theory (one would sum the diagrams of Fig. 4 with the two vertices labeled by x and y). However, in this same phase there is charge screening for large distances, and hence the opposite effect one wants in confinement: Largely separated charges have no interaction with each other; they merely interact individually with the vacuum. In the dipole phase, different effects arise. If one places a number of widely separated charges into the vacuum, the vacuum will immediately produce the opposite charges to form dipoles at the cost of $\beta\mu$ per charge. On this physical basis one can write

$$\langle e^{i\beta q\chi(\rho_1)}\cdots e^{i\beta q\chi(\rho_n)}\rangle\sim\lambda_R{}^n\,, \tag{8.1}$$

where $\int\rho_i(x)d^2x=\pm1$, ρ_i peaked about x_i and all x_i widely spaced. Because charges are immediately turned into dipoles the Green's function $G(x,y)$ [Eq. (7.14)] is constant for large $|x-y|$, and the two charges will not be confined.

Now consider the situation when fractional charges (say a $+\frac{1}{2}$ and a $-\frac{1}{2}$ are placed in the vacuum). These fractional charges cannot form dipoles because the vacuum quanta are integral.

The fractional charges are not screened. The interaction between the $+\frac{1}{2}$ and the $-\frac{1}{2}$ will be essentially a logarithm mitigated by dipole effects. The dipole strength (which is determined by renormalization) will govern how much the dipoles influence the interaction of the fractional charges. For weak dipoles, one expects the logarithmic potential to remain intact. One can write

$$\langle \exp[i\beta \tfrac{1}{2} q\chi(x)] \exp[-i\beta\tfrac{1}{2}q\chi(y)]\rangle \sim |x-y|^{-2\beta q^2(|x-y|)} , \tag{8.2}$$

with $q(r)$ an effective charge satisfying $q(r)\to\frac{1}{2}$ for $r\to 0$ (asymptotic freedom) and $q(r)$ slowly varying for large r. This type of charge screening of the Green's function is reminiscent of a similar phenomenon[11] found in the massive Schwinger model. In general one has

$$\langle e^{i\beta(N+f)q\chi(x)} e^{-i\beta(N+f)q\chi(y)}\rangle \sim \lambda_R{}^{2N}|x-y|^{-2\beta f^2 q^2(|x-y|)} , \tag{8.3}$$

with $0\leq f<1$, fq being the fractional excess charge. For $f=0$ the Green's functions are roughly constant. One can, of course introduce triality operators

$$\Phi_{q/3}(x) = \exp[i\beta\tfrac{1}{3}q\chi(x)]$$

and

$$\Phi^*_{q/3}(x) = \exp[-i\beta\tfrac{1}{3}q\chi(x)] = \Phi_{-q/3}(x) .$$

The Green's functions for these operators vanish unless the number of $\Phi_{q/3}$'s minus the number of $\Phi^*_{q/3}$'s is three times as integer. $\Phi_{q/3}\Phi^*_{q/3}$ configurations and $\Phi_{q/3}\Phi_{q/3}\Phi_{q/3}$ configurations exist. This resembles the triality of the quark model. Of course, there also are n-ality operators $\Phi_{q/n} = e^{i\beta(q/n)\chi(x)}$ and $\Phi^*_{q/n}(x)=\Phi_{-q/n}(x)$. It is the nature of the vacuum that determines these unusual effects.

Fractional quanta have already been used as a possible quark confinement mechanism in four dimensions. The model in mind is the meron gas of Callan, Dashen, and Gross,[24] where the charge in the theory is not the usual charge but the topological charge, and the particles having fractional $\frac{1}{2}$ charge are the merons. The use of fractional quanta is not as unnatural as one might think.

Another interesting effect is that the relevant operators are of the form $\Phi_\rho = e^{i\beta\chi(\rho)}$ $[\chi(\rho) \equiv \int \chi(x)\rho(x)d^2x]$ since these operators produce charge distributions ρ in the Coulomb analog model. These operators are precisely the ones Coleman used to show the equivalence between the sine-Gordon theory and the massive Thirring theory. Being used to perturbation theory, one usually works with the bare vacuum and approximates the interacting fields by free fields in which case the interesting Green's functions are $\langle\chi(x_1)\cdots\chi(x_n)\rangle$. The sine-Gordon theory shows that such simplistic vacuum expectation values may not be the interesting ones. In fact there is no reason why, in a particular theory, the relevant operators are not complicated functions of the fields. This conclusion may be applicable to gauge theories.

A third unusual effect is due to the infrared divergences in the sine-Gordon two-dimensional theory. The normal Feynman rules are invalid. For example, one would conclude from $\mathcal{L}_I = 2\lambda_0\cos\beta q\chi$ that bubble diagrams of order $\lambda_0{}^3$ (such as in Fig. 10) contribute to the vacuum energy. This is incorrect since it violates charge neutrality. Renormalization procedures would be upset if non-neutrality is not maintained since the use of $\ln(|x-y|/a_0)$ for the propagator would make the theory depend on a_0, which it should not. The modification of naive perturbation theory rules by infrared divergences may affect four-dimensional theories such as the popular gauge theories. If such an effect occurs the usual Feynman rules are wrong and may upset the renormalization of infrared singularities.

Fourth, it is curious that to every finite order (in βq^2) in perturbation theory there are no ultraviolet singularities; yet when all orders are summed an ultraviolet divergence arises when $\beta q^2 \geq 1$. The reader has already seen this in the bubble diagrams of Fig. 4. They have no ultraviolet singularities to any finite order; yet when summed they are $\sim\lambda_R{}^2\int r^{-2\beta q^2}d^2r$ which diverges for $\beta q^2\geq 1$.

IX. SUMMARY

Here is a list of the main results:

A (Sec. II). Certain field theories are equivalent to gases of interacting particles, in particular the following:

1 The sine-Gordon theory corresponds to a neutral Coulomb gas.

2 The "massive" sine-Gordon theory and the massive Schwinger model at zero Coleman angle correspond to a gas of quanta interacting via Yukawa potentials.

B (Sec. III). The Feynman diagrams for these theories have a statistical-mechanical interpretation. The correspondence is outlined in Table I.

C (Sec. V). The vacuum of the two-dimensional sine-Gordon theory undergoes a phase transition

FIG. 10. An infared-forbidden diagram.

at βq^2 near 2. For $\beta q^2 < 2$ there is a plasma phase and for $\beta q^2 > 2$ there is a dipole gas phase. $\beta q^2 = 2$ is precisely the value Coleman[13] finds a vacuum instability.

D (Sec. VI). The two-dimensional nonlinear O(2) σ model with linear symmetry-breaking term is equivalent to the sine-Gordon theory. Results C and D imply the following:

1 The σ model undergoes a phase transition for $f_{\Pi} \approx (1/8\pi)^{1/2}$.

2 The σ model contains solitons and fermions.

E (Sec. VII). When $\beta q^2 < 1$ the sine-Gordon theory [massive Thirring and O(2) σ models] are renormalizable to all orders, i.e., they are well-defined theories.[25]

F (Sec. VIII). The dipole phase of the sine-Gordon theory completely shields integral charges but is unable to do so for fractional charges.

G The relevant operators for the sine-Gordon theory are not simply polynomials in the fields. They are $\Phi_q(x) = e^{i\beta q \chi(x)}$ and have the simple physical interpretation of producing a charge, q, at x. They are the operators used by Coleman to prove the equivalence of the sine-Gordon and massive Thirring models.

H (Sec. VIII). Operators exhibiting the quarklike triality condition are $\Phi_{q/3}(x)$. They have this property because of the infrared singular nature of the sine-Gordon field theory.

I (Secs. IV and VIII). Naive Feynman rules may be incorrect when infrared singularities occur. The sine-Gordon theory exhibits such a property.

J (Secs. VII and VIII). The sine-Gordon theory has no ultraviolet singularities to every finite order of perturbation theory; yet when diagrams are summed an ultraviolet divergence appears.

K Callan, Dashen, and Gross[26] have recently shown that the instanton approximation to two-dimensional charged scalar electrodynamics with massless fermions is equivalent to a neutral Coulomb gas. The instantons are Nielsen-Olesen vortices. The effect of the massless fermions is to raise the inverse temperature βq^2 from 0 to N, the number of fermions. Since the Coulomb interaction is mitigated at short distances their model has a natural renormalization. *Their result together with result I of this paper implies their model will possible have (sine-Gordon) solitons.* The existence of such solitons depends on how much the Coulomb force is modified and how good the instanton approximation is.

There is a good chance that a theory of strong interactions in four dimensions will have many of the properties exhibited by the two-dimensional sine-Gordon theory. It is conceivable that the hadron vacuum has a complicated structure which must be treated using statistical mechanics. Strange effects can occur when such a vacuum has "a lot of quanta" in it. It would be able to support asymptotic freedom because at short distances the quanta are ineffective and at the same time it could provide confinement since at large distances the many-body effects of such quanta can be unusual. It is conceivable that stringlike structures or other types of extended objects could condense out of the vacuum when other quanta such as quarks are introduced. Further unusual effects created by a strong-interaction vacuum may provide for the triality condition now observed.

ACKNOWLEDGMENTS

I would like to thank Professor Bardakci for interesting discussions and for reading the manuscript. This research was sponsored by the National Science Foundation under Grant No. PHY77-23512.

---

[1]See, for example, R. Brout, Phys. Rep. 10C, 1 (1974).

[2]For a review see F. W. Wiegel, Phys. Rep. 16C, 57 (1975).

[3]S. Albeverio and R. Høegh-Krohn, Commun. Math. Phys. 30, 171 (1973); J. Fröhlich, Phys. Rev. Lett. 34, 833 (1975); J. Fröhlich, in *Renormalization Theory*, edited by G. Velo and A. S. Wightman (Reidel, Boston, 1976); D. Brydges, Commun. Math. Phys. 58, 313 (1978).

[4]J. Fröhlich and E. Seiler, Helv. Phys. Acta 49, 889 (1976); J. Fröhlich and Y. M. Park, *ibid.* 50, 315 (1977).

[5]J. Fröhlich, Commun. Math. Phys. 47, 233 (1976).

[6]A. M. Polyakov, Nucl. Phys. B120, 429 (1977).

[7]See, for example, J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975), or L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon, London, 1970).

[8]This is due to boson-fermion correspondence in two dimensions. For a summary see M. Bander, Phys. Rev. D 13, 1566 (1976) and the references therein.

[9]For two dimensions (which will be needed later) the propagator is $(2/\beta)\ \ln\ (|r|/a_0)$ where a_0 is an arbitrary scale.

[10]See, for example, A. Isihara, *Statistical Physics* (Academic, New York, 1971).

[11]The Schwinger model (both massive and massless) has many interesting properties. See J. Lowenstein and A. Swieca, Ann. Phys. (N. Y.) 68, 172 (1971); A. Casher, J. Kogut, and L. Susskind, Phys. Rev. D

10, 732 (1974); S. Coleman, R. Jackiw, and L. Susskind, Ann. Phys. (N. Y.) 93, 267 (1975). For a review of this last work see L. Susskind and J. Kogut, Phys. Rep. 23C, 348 (1976).

[12]There is an enormous amount of literature on the sine-Gordon theory. A good but limited set of references on the field-theory aspects is contained in Fröhlich's Erice lectures (Ref. 3), while a good but again limited set of references to the classical and solitons aspects is contained in Ref. 18.

[13]S. Coleman, Phys. Rev. D 11, 2088 (1975).

[14]M. B. Halpern and P. Senjanović, Phys. Rev. D 15, 3629 (1977).

[15]E. H. Hauge and P. C. Hemmer, Phys. Norvegica 5, 209 (1971); J. M. Kosterlitz and D. J. Thouless, J. Phys. C 6, 1181 (1973); J. M. Kosterlitz, *ibid.* 7, 1046 (1974); C. Deutsch and M. Lavard, Phys. Rev. A 9, 2598 (1974).

[16]S. Mandelstam, Phys. Rev. D 11, 3026 (1975).

[17]A. Luther, Phys. Rev. B 14, 2153 (1976). Mass ratios were first obtained using semiclassical techniques. See R. F. Dashen, B. Hasslacher, and A. Neveu, Phys. Rev. D 11, 3424 (1975). A review of this work is presented by R. Rajaraman, Phys. Rep. 21C, 227 (1975).

[18]A. C. Scott, F. Y. F. Chu, and D. W. McLaughlin, Proc. IEEE 61, 1443 (1973).

[19]I would like to thank K. Bardakci for this point.

[20]I thank K. Bardakci for showing me this.

[21]E. Brézin and J. Zinn-Justin, Phys. Rev. Lett. 36, 691 (1976); Phys. Rev. B 14, 3110 (1976).

[22]W. A. Bardeen, B. W. Lee, and R. E. Shrock, Phys. Rev. D 14, 985 (1976).

[23]For similar results on the convergence of this theory see J. Fröhlich's work in Refs. 3, 4, and 5.

[24]C. G. Callan, R. Dashen, and D. J. Gross, Phys. Lett. 66B, 375 (1977).

[25]J. Fröhlich has proved that in such a case ϕ is a rigorously defined Wightman field. See Refs. 3 and 4.

[26]C. G. Callan, R. Dashen, and D. J. Gross, Phys. Rev. D 16, 2526 (1977).

FIG. 1. Some Feynman graph combinatorial factors. In (a) there are four lines connecting the two vertices. According to rule (e) there is a factor of $1/4!$. (b) is a threefold self-energy tadpole. According to rule (f) there is a factor of $1/3!\,2^3$.

FIG. 10. An infared-forbidden diagram.

FIG. 2. The effect of self-energy tadpoles. The bare vertex is replaced by a sum of terms, each one with an additional tadpole attached.

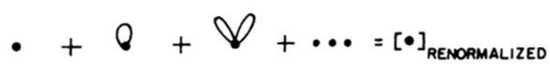

FIG. 3. The diagrams of order λ_0.

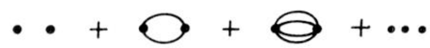

FIG. 4. The diagrams of order λ_R^2. The self-energy tadpoles are dropped according to the modified (b) and (f) rules.

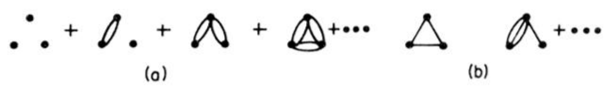

FIG. 5. Diagrams of order λ_R^3. Diagrams of (a) have an even number of propagators between vertices, whereas (b) have only an odd number.

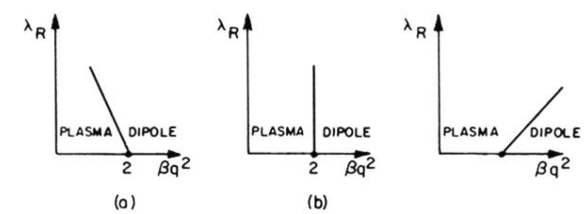

FIG. 6. Three possible phase diagrams.

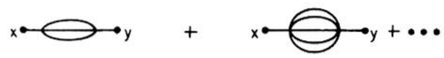

FIG. 7. Contributions to the two-point function. Although individual graphs are not divergent, the sum of these graphs gives a divergent contribution to the two-point function when $\beta q^2 \geq 1$.

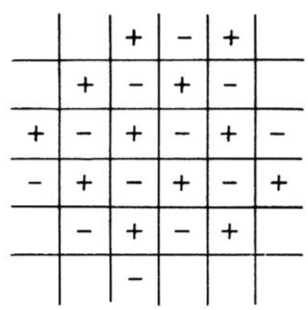

FIG. 8. Minimum-energy configuration.

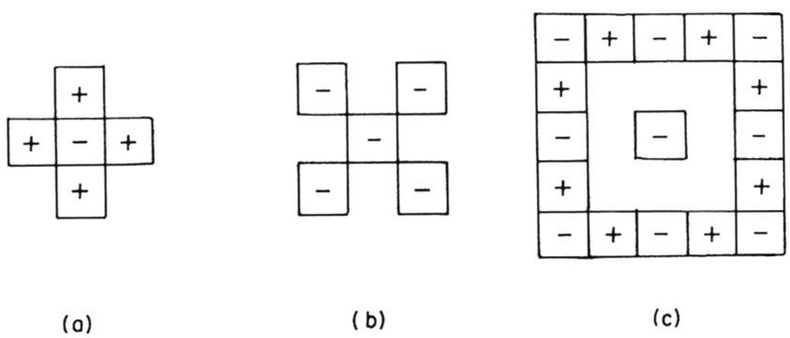

FIG. 9. Leading contributions to the energy of a minus charge.