Localization transition on the Bethe lattice

Martin R. Zirnbauer
W. K. Kellogg Radiation Laboratory, California Institute of Technology, Pasadena, California 91125
(Received 21 March 1986)

The critical behavior of disordered single-particle systems without time-reversal invariance is described by a nonlinear σ model formulated in terms of graded pseudounitary matrices. It is shown that the symmetry-breaking pattern of this model predicts the correct analytical structure of the two-point Green's functions of the disordered system, which are expected to be singular (finite) for localized (extended) states in the limit of vanishing frequency. The main purpose of the paper is to obtain exact solutions for the graded nonlinear σ model on a Bethe lattice. A combination of analytical and numerical methods is used to determine the critical behavior of all two-point Green's functions. In contrast with other work on the graded nonlinear σ model, no minimum metallic conductivity is found. Instead, the averaged inverse conductivity has an exponential singularity at the critical point.

I. INTRODUCTION

It is now quite generally accepted that disordered single-particle systems in more than two space dimensions may be in either one of two distinct phases. For strong disorder, or energies far away from the band center, the eigenfunctions of the random Hamiltonian are localized, while for weak disorder, or energies close to the band center, the eigenfunctions extend over the whole system. What is still in dispute is what happens at the critical point where the phase transition from localized to extended states occurs. Wegner[1] first conjectured that the critical behavior of (time-reversal noninvariant) disordered single-particle systems may be described by a nonlinear σ model formulated in terms of $2n\times 2n$ matrices Q [$Q\in \mathrm{U}(n,n)/\mathrm{U}(n)\otimes\mathrm{U}(n)$], with $n=0$ due to the replica trick. This conjecture made available the perturbative renormalization-group equations worked out for closely related models[2] and gave predictions for critical exponents close to two dimensions. In addition to Wegner's noncompact model, a corresponding compact parametrization has also been in use.[3-6] It is not clear whether the predictions of these models can be made to agree in the limit of a vanishing number of replicas, although this seems unlikely in general.

In a more recent line of development, Efetov has introduced[7] a related nonlinear σ model, which avoids the use of the replica trick altogether, at the cost of being defined over a *graded* coset space. This model will be referred to here as the graded pseudounitary nonlinear σ model. It is my personal conviction, not widely shared at the present moment, that among the various nonlinear σ models proposed, Efetov's model is the only one that is capable of providing a true description of the localization transition. This conviction is based on experience with some soluble cases: The case of a single large random matrix[7,8] and the analysis of the model on a Cayley tree (Bethe lattice). Further evidence comes from the observation that the symmetry-breaking pattern for the graded model predicts the correct analytical structure for the Green's functions of the disordered system. This cannot be said of either of the nonlinear σ models formulated in terms of $2n\times 2n$ matrices, $n=0$. Even from a purely technical point of view, much is gained by making the initial investment of formulating the localization problem with the use of graded algebras. I believe that it will be easier in the long term to work with 4×4 (or 8×8) graded matrices than to solve the replica field theory for *all* values of n and perform the analytic continuation $n\to 0$.

Although the analysis of the graded pseudounitary nonlinear σ model is still in its infancy, several facts are known at this point. Its perturbation expansion agrees with what is obtained from replica field theory, which means that the ϵ-expansion results ($\epsilon=d-2$) remain unchanged. New insight is now emerging from the analysis of the model on the Bethe lattice. This work was again initiated by Efetov,[9,10] who claimed, however, to find some very startling results for the phase transition of the model ("mobility transition") which seem difficult to reconcile with the scaling theory of localization.[11] (In a very recent letter,[12] Efetov has even gone so far as to challenge the notion of universality and the existence of a renormalization group.) Rather surprisingly, Efetov's analysis provides neither a complete nor, in fact, a totally correct picture of the mobility transition on the Bethe lattice. It is the main purpose of this paper to clarify Efetov's results and to complete his analysis. For reasons of technical simplicity, I will restrict myself to the case of unitary symmetry, i.e., to systems without time-reversal invariance.

The results of this paper are summarized as follows. On the Bethe lattice, the two-point Green's functions $K^{(1)}$ and $K^{(2)}$ [defined in Eqs. (3.1)] of the graded pseudounitary nonlinear σ model behave as

$$K^{(1)}(x,y;\gamma,\epsilon)\simeq\frac{1}{\epsilon}\Gamma^{\mathrm{loc}}(x-y;\gamma)+O(\epsilon^0)$$
$$\simeq K^{(2)}(x,y;\gamma,\epsilon) \tag{1.1}$$

(ϵ infinitesimal) for $\gamma<\gamma_c$ (localized states). $\Gamma^{\mathrm{loc}}(0;\gamma)$

reaches a minimum value as $\gamma \to \gamma_c$ and goes discontinuously to zero at the critical point $\gamma = \gamma_c$. $\Gamma^{\rm loc}(x;\gamma)$ for $|x| \gg 1$ has the critical behavior.

$$m^{|x|}\Gamma^{\rm loc}(x;\gamma) \simeq c\,|x|^{-3/2}\exp(-|x|/\xi)\,, \tag{1.2}$$

where the inverse correlation length ξ^{-1} vanishes linearly, $\xi^{-1} \propto |\gamma - \gamma_c|$, and $m+1$ is the coordination number of the lattice. These results are consistent with a sum rule for $\Gamma^{\rm loc}$, Eq. (4.12). On the side of extended states $(\gamma > \gamma_c)$,

$$K^{(1)}(x,y;\gamma,\epsilon) \simeq 1 + \Gamma^{\rm ext}(x-y;\gamma) + O(\epsilon) \tag{1.3}$$

and $\Gamma^{\rm ext}$ exhibits a strong singularity near the critical point:

$$\Gamma^{\rm ext}(0;\gamma) \simeq A\exp\{+[B(\gamma-\gamma_c)]^{-1/2}\}\,. \tag{1.4}$$

The correlation function $\Gamma^{\rm ext}(x;\gamma)$ again follows the law (1.2) but with a different correlation length ξ and only in the range $1 \ll |x| \ll \xi$. For $|x| \gg \xi$ the decay is purely exponential,

$$m^{|x|}\Gamma^{\rm ext}(x;\gamma) \propto \exp(-|x|/\xi)\,. \tag{1.5}$$

The two-point Green's function $K^{(2)}$ is more difficult to control but behaves like $K^{(1)}$ as long as $|x| \ll \xi$. For $|x| \gg \xi$ and $\epsilon \to 0$, this function becomes qualitatively more long ranged.

The present paper is organized in the following way. In Sec. II the basic definitions for the graded pseudounitary nonlinear σ model are reproduced. The connection with localization theory imposes several constraints on the Green's functions of this model. Some of these are reviewed in Sec. III. The symmetry-breaking pattern of the model is then discussed, and an order parameter is introduced, in Sec. IV. The bulk of the paper is contained in Sec. V, where the critical behavior of the two-point Green's functions is analyzed for both localized and extended states on the Bethe lattice. My concluding remarks are presented in Sec. VI.

It should be mentioned here that localization on the Bethe lattice has previously been studied within the framework of the original Anderson model by Abou-Chacra, Anderson, and Thouless[13] and by Kunz and Souillard.[14] I feel that it would be of considerable interest to try and compare the results obtained in the present paper with this earlier work. Unfortunately, there are several differences that hamper such a comparison. First, the nonlinear σ model studied here has a unitary symmetry, appropriate for disordered single-particle systems without time-reversal invariance, while Anderson's model applies to time-reversal invariant systems. Unbroken time-reversal invariance requires the study of an analogous nonlinear σ model with orthogonal symmetry (or rather, to be precise, orthosymplectic symmetry). It is not clear whether the orthosymplectic and unitary models are equivalent on the Bethe lattice. Second, the nonlinear sigma model considered here derives from Wegner's N-orbital model in the limit of large N, whereas in Anderson's model $N=1$. Third, the emphasis in the present paper is on the determination of critical exponents. In the classic paper of Abou-Chacra *et al.* exponents were not yet obtained. The exponent given by Kunz and Souillard cannot be evaluated directly in the nonlinear σ model because it describes a Green's function that involves an absolute value, $|\Psi(x)\Psi(0)|$, of the eigenfunctions. For these reasons a comparison with the two quoted papers will not be attempted here.

II. NONLINEAR σ MODEL WITH GRADED SYMMETRY

A complete definition of the graded pseudounitary nonlinear σ model was given in Sec. 3 of Ref. 15. To set the stage for what is to come, I will reproduce some of the most important definitions; for quantities not defined here, I refer to this reference.

The fields of the model are 4×4 matrices, denoted by Q, which are elements of a graded coset space. They can be written as $Q = -iT\sigma_3 T^{-1}$, with $T \in {\rm U}(1,1/2)$ a 4×4 graded pseudounitary matrix. Technically speaking, $Q \in {\rm U}(1,1/2)/{\rm U}(1/1)\otimes{\rm U}(1/1)$. A particularly useful parametrization introduced by Efetov[7] is

$$Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} = U \begin{pmatrix} -i\Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & +i\Lambda_{22} \end{pmatrix} U^{-1}, \quad U = \begin{pmatrix} u & 0 \\ 0 & v \end{pmatrix}, \tag{2.1a}$$

$$\Lambda_{11} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \Lambda_{22}, \quad \Lambda_{12} = \begin{pmatrix} \mu_1 & 0 \\ 0 & i\mu_2^* \end{pmatrix}, \quad \Lambda_{21} = \begin{pmatrix} \mu_1^* & 0 \\ 0 & i\mu_2 \end{pmatrix}. \tag{2.1b}$$

Here, u and v are graded unitary $[u,v \in {\rm U}(1/1)]$, and the "eigenvalues" λ_1,λ_2, $\mu_1 = |\mu_1| e^{i\phi_1}$, $\mu_2 = |\mu_2| e^{i\phi_2}$ obey the constraints

$$\lambda_1^2 - |\mu_1|^2 = 1, \quad \lambda_2^2 + |\mu_2|^2 = 1\,, \tag{2.2}$$

which is equivalent to requiring that $Q^2 \equiv -1$. Equation (2.2) states that the vectors $(\lambda_1, {\rm Re}\mu_1, {\rm Im}\mu_1)$ and $(\lambda_2, {\rm Re}\mu_2, {\rm Im}\mu_2)$ are situated on the coset spaces $H_2 \simeq {\rm SU}(1,1)/{\rm U}(1)$ and $S_2 \simeq {\rm SU}(2)/{\rm U}(1)$, respectively.

The action of the model consists of two terms. The first term, $S_0[Q]$, has an invariance under global transformations $Q(x) \to TQ(x)T^{-1}$, $T \in {\rm U}(1,1/2)$. It represents the usual interaction term of the nonlinear σ model and on the lattice takes the form

$$S_0[Q] = -\frac{\gamma}{2} \sum_{(x_1,x_2)} {\rm trg} Q(x_1)Q(x_2)\,, \tag{2.3}$$

where the sum runs over pairs of neighboring lattice sites. γ^{-1} plays the role of a coupling constant. The second term, $S_1[Q]$, breaks the global U(1,1/2) symmetry down to ${\rm U}(1/1)\otimes{\rm U}(1/1)$, i.e., to global transformations $Q(x) \to UQ(x)U^{-1}$ with U as in Eq. (2.1a). This term has the explicit form

$$S_1[Q] = i\epsilon \sum_x {\rm trg}\sigma_3 Q(x)\,. \tag{2.4}$$

The dimensionless quantity ϵ corresponds to the imaginary part of the energy, η, that appears in the argument of the Green's function for the disordered single-particle

system. It will be important that the global $U(1/1)\otimes U(1/1)$ invariance, here referred to as "graded symmetry," is an *exact* and hence unbroken symmetry of the model.

Expectation values of fields are defined as usual:

$$\langle A(Q)\rangle = \frac{1}{Z}\int d[Q]A(Q)e^{-S[Q]},$$

$$S[Q]=S_0[Q]+S_1[Q]. \quad (2.5)$$

The normalization factor $Z=\int d[Q]\exp(-S[Q])$ can be omitted because $Z\equiv 1$, due to the exact graded symmetry.

As was mentioned in the Introduction, the nonlinear σ model studied here is derived most directly from the N-orbital model of Wegner.[16] A detailed exposition of the mathematical formalism required for this derivation was given in Ref. 17. This material will not be repeated here. However, it may be useful to write down the explicit relations connecting the parameters γ and ϵ to some actual random potential. Let us consider Wegner's N-orbital model on a d-dimensional cubic lattice with site-diagonal and nearest-neighbor hopping matrix elements, which are taken to be Gaussian distributed random variables with zero mean and variance $\lambda^2(1-2dr)$ and $\lambda^2 r$, respectively. For large N, we then have

$$\gamma=2Nr(\pi\lambda\bar{\rho})^2,\quad \epsilon=\pi N\eta\bar{\rho} \quad (2.6)$$

where

$$\bar{\rho}(E)=\frac{1}{\pi\lambda}[1-(E/2\lambda)^2]^{1/2}. \quad (2.7)$$

These relations result from making a saddle-point approximation and subsequently equating the interaction strengths of the long-wavelength (small-momentum) modes of the N-orbital model and of the nonlinear σ model, without taking into account the renormalization caused by large momentum components.

III. TWO-POINT GREEN'S FUNCTIONS

The basic hypothesis underlying this work is that the critical behavior of the graded pseudounitary nonlinear σ model and the critical behavior of disordered single-particle systems without time-reversal symmetry are identical. In particular, both systems are assumed to have the same phase structure. Given the knowledge that we have about disordered single-particle systems, this hypothesis results in certain constraints (or expectations) for the Green's functions of the model defined in Eqs. (2.1)—(2.5). Several of these constraints will be listed below, and it will then be explained why they fit very naturally into the symmetry-breaking pattern for the σ model. The results for the Bethe lattice derived in Sec. V will also give support to the above hypothesis.

There exist only two nontrivial two-point Green's functions for the present model:[15]

$$K^{(1)}(x,y;\gamma,\epsilon)=\langle Q^{ll}_{11}(x)Q^{kk}_{22}(y)\rangle = \tfrac{1}{4}\langle \mathrm{trg}kQ_{11}(x)\mathrm{trg}kQ_{22}(y)\rangle, \quad (3.1a)$$

$$K^{(2)}(x,y;\gamma,\epsilon)=\langle Q^{lk}_{12}(x)(-1)^{[k]}Q^{kl}_{21}(y)\rangle = \tfrac{1}{4}\langle \mathrm{trg}kQ_{12}(x)kQ_{21}(y)\rangle. \quad (3.1b)$$

The two phases of the nonlinear σ model can be distinguished by studying these Green's functions in the limit of vanishing ϵ. For γ less than some critical value, γ_c, the $K^{(n)}(x,y;\gamma,\epsilon)$ ($n=1,2$; x,y arbitrary) are expected to diverge as ϵ^{-1}. This defines the "disordered" or "high-temperature" phase associated with localized states. On the other hand, for $\gamma>\gamma_c$ all two-point Green's functions should attain a definite finite value when $\epsilon\to 0$. In this case the system is in the "ordered" or "low-temperature" phase that corresponds to extended states. These constraints on the two-point Green's functions follow from quite elementary considerations for the disordered single-particle system; see Ref. 18.

The expected behavior of the $K^{(n)}$ invites the definition of special correlation functions in the limit $\epsilon\to 0$. Two quantities frequently used in the following are

$$\Gamma^{\mathrm{loc}}(x-y;\gamma)=\lim_{\epsilon\to 0}\epsilon K^{(1)}(x,y;\gamma,\epsilon), \quad (3.2a)$$

$$\Gamma^{\mathrm{ext}}(x-y;\gamma)=\lim_{\epsilon\to 0}[\langle Q^{ll}_{11}(x)Q^{kk}_{22}(y)\rangle - \langle Q^{ll}_{11}(x)\rangle\langle Q^{kk}_{22}(y)\rangle] = \lim_{\epsilon\to 0}[K^{(1)}(x,y;\gamma,\epsilon)-1]. \quad (3.2b)$$

Γ^{loc} and Γ^{ext} will be referred to as the density-density correlation function for localized and extended states, respectively.

It is perhaps worth reviewing why the one-point Green's function of the graded nonlinear σ model, $\langle Q^{kk}_{pp}(x)\rangle$, is noncritical and thus without interest. The presence of the matrix element $Q^{kk}_{pp}(x)$ destroys the graded symmetry of the integrand, but only imperfectly so: one of the graded unitary subgroups in $U(1/1)\otimes U(1/1)$ is broken but the other one is left intact. Due to this residual graded symmetry, the functional integral $\langle Q\rangle$ can be evaluated trivially, and the result is that it is simply given by the value of the integrand at the origin of the coset space, $Q=-i\sigma_3$:

$$\langle Q^{kk}_{pp}(x)\rangle=(-i\sigma_3)^{kk}_{pp}e^{-S[-i\sigma_3]}=-i(-1)^{p-1}, \quad (3.3)$$

which is *independent* of γ and ϵ. [The origin of the coset space will sometimes be denoted by $Q=-i\sigma_3$ and sometimes by $(\lambda_1,\lambda_2)=(1,1)$.] By the same reasoning, two-point functions that involve only matrix elements of Q in the 11 block, say, are trivial. For example, $\langle Q^{ll}_{11}(x)Q^{kk}_{11}(y)\rangle\equiv -1$.

IV. SYMMETRY BREAKING AND ORDER PARAMETER

The symmetry-breaking pattern for the model (2.1)—(2.5) is clearly laid out by the definitions given in Sec. II. There will be a disordered phase (strong coupling, $\gamma<\gamma_c$) in which the global U(1,1/2) symmetry is restored as the symmetry-breaking term $S_1[Q]$ is removed. There will, in general, also be an ordered phase (weak coupling,

$\gamma > \gamma_c$), where this symmetry is spontaneously broken; that is to say the symmetry is not recovered in the limit $\epsilon \to 0$. In the following, I will make these statements more concise, and I will indicate why they are consistent with all that is required of a symmetry-breaking pattern which is to describe the localization transition.

For a given, infinitely extended lattice in d dimensions, let us pick an arbitrary site x and integrate e^{-S} over all matrices except the one belonging to that site, $Q = Q(x)$. The result of this integration is a function of Q, denoted by $F(Q)$. Because graded symmetry is exact, $F(Q)$ depends only on the two (independent) graded invariants that can be constructed from the matrix elements of Q. These are the "eigenvalues" λ_1 and λ_2, see Eq. (2.1b):

$$F(Q) = F(UQU^{-1}) = F(\lambda_1, \lambda_2) . \tag{4.1}$$

Given the function F, we can calculate the one-point Green's function, and the diagonal part of the two-point Green's functions:

$$\langle Q_{pp}^{kk}(x) \rangle = \int d\mu(Q) Q_{pp}^{kk} F(Q) , \tag{4.2a}$$

$$K^{(1)}(x,x;\gamma,\epsilon) = \int d\mu(Q) Q_{11}^{ll} Q_{22}^{kk} F(Q) , \tag{4.2b}$$

$$K^{(2)}(x,x;\gamma,\epsilon) = \int d\mu(Q) Q_{12}^{lk} (-1)^{[k]} Q_{21}^{kl} F(Q) . \tag{4.2c}$$

The integrals over the graded matrix Q can be reduced to ordinary real integrals by using an integral theorem proven in Ref. 15. This theorem states that the integral of any function $f(Q)$ (with analytical properties as defined in Ref. 15) is given by

$$\int d\mu(Q) f(Q) = f(-i\sigma_3) + \int_1^\infty d\lambda_1 \int_{-1}^{+1} d\lambda_2 (\lambda_1 - \lambda_2)^{-2} f_4(\lambda_1, \lambda_2) . \tag{4.3}$$

Here, f_4 is the coefficient of the highest term in the Taylor series expansion of $f(Q)$ with respect to the anticommuting variables parametrizing the matrices u and v in Eq. (2.1a). (If this coefficient has a dependence on the angular variables ϕ_1 and ϕ_2 in addition to that on λ_1 and λ_2, then an angular average is to be taken; see Ref. 15.) Application of this theorem to the integrals (4.2) yields[10,15]

$$\langle Q_{pp}^{kk}(x) \rangle = -i(-1)^{p-1} F(-i\sigma_3) , \tag{4.4a}$$

$$K^{(1)}(x,x;\gamma,\epsilon) = 1 + \int_1^\infty d\lambda_1 \int_{-1}^{+1} d\lambda_2 F(\lambda_1, \lambda_2) , \tag{4.4b}$$

$$K^{(2)}(x,x;\gamma,\epsilon) = \int_1^\infty d\lambda_1 \int_{-1}^{+1} d\lambda_2 \frac{\lambda_1 + \lambda_2}{\lambda_1 - \lambda_2} F(\lambda_1, \lambda_2) , \tag{4.4c}$$

where in the second equation I have used that $F(-i\sigma_3) = F(1,1) = 1$, another direct consequence of the exact graded symmetry; see Eq. (3.3) and compare to Eq. (4.4a).

I shall now argue that $F(Q)$ is a convenient order parameter that can be used to label the two phases of the system. Let us first discuss what happens in the "disordered" phase where the U(1,1/2) symmetry is restored as $\epsilon \to 0$. Symmetry restoration implies that F becomes invariant under U(1,1/2) in this limit. The only U(1,1/2) invariant that can be constructed from a single matrix Q is the trivial one, $F = \text{const}$. Formally, this is expressed by

$$\lim_{\epsilon \to 0} \lim_{V \to \infty} F(\lambda_1, \lambda_2; \gamma, \epsilon) = 1 \quad (\gamma < \gamma_c) , \tag{4.5}$$

where $V \to \infty$ denotes the thermodynamic limit, and the parametric dependence of F on γ and ϵ has been included in the list of its arguments. By detailed inspection of the integrand and counting powers of ϵ, I was able to show in Sec. 6 of Ref. 15 that the leading behavior of F (and of other functions) in the limit $\epsilon \to 0$ can be extracted by absorbing one factor of ϵ into each of the integration variables. In the present context this means that F can be written as

$$F(\lambda_1, \lambda_2) = f(\epsilon\lambda_1, \epsilon\lambda_2) + O(\epsilon) . \tag{4.6}$$

Assuming a definite form for the function f, we see that the convergence of F to unity as $\epsilon \to 0$ is *uniform* in λ_2, but *pointwise* in λ_1. Because the latter integration variable is noncompact, symmetry restoration causes the two-point functions to be singular at $\epsilon = 0$. The order of the singularity is obtained by inserting (4.6) into Eqs. (4.4b) and (4.4c):

$$K^{(n)}(x,x) \simeq \int_1^\infty d\lambda_1 \int_{-1}^{+1} d\lambda_2 f(\epsilon\lambda_1, \epsilon\lambda_2) \simeq \frac{2}{\epsilon} \int_0^\infty d\omega f(\omega, 0) \propto \frac{1}{\epsilon} , \tag{4.7}$$

which is precisely what is required for localized states.

On the other hand, for sufficiently weak coupling and dimension $d > 2$, the global U(1,1/2) symmetry is expected to be spontaneously broken. That is,

$$\lim_{\epsilon \to 0} \lim_{V \to \infty} F(\lambda_1, \lambda_2; \gamma, \epsilon) = \bar{F}(\lambda_1, \lambda_2; \gamma) \quad (\gamma > \gamma_c) , \tag{4.8}$$

with $\bar{F}$ a function that decays as λ_1 and λ_2 depart from the origin of the coset space, $(\lambda_1, \lambda_2) = (1,1)$. $\bar{F}$ will actually have the (somewhat stronger) property of being normalizable; this is what is needed to make the $K^{(n)}(x,x;\gamma,\epsilon)$ finite at $\epsilon = 0$, which is the proper behavior for the two-point Green's functions in the regime of extended states.

Assuming henceforth the equivalence between localization and the graded nonlinear sigma model, I will frequently use (and have already used) the term "localized states" when what I really mean is the phase of unbroken symmetry. Similarly, the word "extended states" will be used as synonymous with the phase of broken symmetry.

The remainder of this section is devoted to the two-point functions $K^{(n)}(x,y)$ for $x \neq y$. These can be calculated from a generalization of the above function, $F(Q,Q')$, obtained by integrating e^{-S} over all matrices except two,

namely $Q=Q(x)$ and $Q'=Q(y)$. Just as in the case $x=y$, the symmetry-breaking pattern implies that $K^{(n)}(x,y;\gamma,\epsilon)\propto\epsilon^{-1}(\epsilon^0)$ for $\gamma<\gamma_c(\gamma>\gamma_c)$, ϵ infinitesimal. A less trivial result is that

$$\lim_{\epsilon\to 0}\epsilon K^{(1)}(x,y;\gamma,\epsilon)=\lim_{\epsilon\to 0}\epsilon K^{(2)}(x,y;\gamma,\epsilon) \tag{4.9}$$

for localized states. A brief sketch of the proof is as follows. (I strictly adhere to the notation defined in Ref. 15.) (i) Study the transformation behavior of the matrices Q and Q' under $\mathrm{U}(1/1)\otimes\mathrm{U}(1/1)$. (ii) Use this to find all $\mathrm{U}(1/1)\otimes\mathrm{U}(1/1)$ invariants that can be constructed from Q and Q'. Some examples are found in Eq. (5.46). (iii) Perform the rescaling of variables appropriate to the limit $\epsilon\to 0$ for localized states [see Eq. (6.20) of Ref. 15]. (iv) Expand $F(Q,Q')$ with respect to the nilpotent terms $\Delta\hat{\alpha}^*$, $\Delta\hat{\alpha}$, $\Delta\hat{\beta}^*$, and $\Delta\hat{\beta}$. (v) Combine the result of this expansion with Eq. (5.15) of Ref. 15 and a similar expression for $K^{(2)}$ and count powers of ϵ to show that only the term proportional to $\Delta\hat{\alpha}^*\Delta\hat{\alpha}\Delta\hat{\beta}^*\Delta\hat{\beta}$ survives in the limit $\epsilon\to 0$. The relation (4.9) then follows. This is a very pleasing result because the underlying disordered single-particle system from which the nonlinear σ model is derived has a corresponding property:

$$\lim_{\eta\to 0}\eta\left[\left\langle x\left|\frac{1}{E-i\eta-H}\right|x\right\rangle\left\langle y\left|\frac{1}{E+i\eta-H}\right|y\right\rangle\right]_{\mathrm{av}}$$
$$=\lim_{\eta\to 0}\eta\left[\left\langle x\left|\frac{1}{E-i\eta-H}\right|y\right\rangle\left\langle y\left|\frac{1}{E+i\eta-H}\right|x\right\rangle\right]_{\mathrm{av}}. \tag{4.10}$$

($[\ldots]_{\mathrm{av}}$ denotes the average over the disorder.) The global U(1,1/2) symmetry of $S_0[Q]$ results in a Ward identity for $K^{(2)}$:

$$\epsilon\sum_y\langle Q_{12}^{lk}(x)(-1)^{[k]}Q_{21}^{kl}(y)\rangle=\frac{i}{2}[\langle Q_{11}^{ll}(x)\rangle-\langle Q_{22}^{kk}(x)\rangle]\,. \tag{4.11}$$

When combined with this identity and the triviality of the one-point Green's function, Eq. (3.3), the relation (4.9) yields a sum rule for Γ^{loc}:

$$\sum_y\Gamma^{\mathrm{loc}}(x-y;\gamma)\equiv 1\quad(\gamma<\gamma_c)\,. \tag{4.12}$$

This constraint will be important in Sec. V A.

V. BETHE LATTICE

Although the model (2.1)–(2.5) has the general form of a nonlinear σ model, attempts at its exact or approximate solution face difficulties that are even more severe than is usually the case with models in statistical mechanics. The model differs from the standard nonlinear σ model in two major respects. (i) The appearance of Grassmann variables makes the model intractable by numerical methods, at least in the form (2.1)–(2.5); it is also unclear whether mean-field techniques can be applied. (ii) The noncompact symmetry of the bosonic-bosonic variables has the consequence that the various observables are nonanalytic at $\gamma=0$; this rules out the use of standard strong-coupling expansions in the analysis of (2.5).

For the case of localized states, it was shown in Ref. 15 how to eliminate the Grassmann variables and formulate a Monte Carlo algorithm that is (or appears to be) computationally feasible. It is not advisable, however, to venture into large-scale numerical simulations without seeking a good qualitative understanding of the model first. It was partly for this reason that I turned my attention to the analysis of the graded pseudounitary nonlinear σ model on the Bethe lattice (also referred to as the "Cayley tree"). Simplifications due to the absence of closed loops on the Bethe lattice make this model tractable by analytical means, and it should be possible to work out the nature of the mobility transition for this case. Results obtained in Ref. 15 agreed with a preliminary account of Efetov's work[9] to the extent that the results overlapped. However, as explained in the Introduction, recent further publications by Efetov[10,12] have motivated me to take another look at the problem.

On the Bethe lattice, evaluation of the functional integral (2.5) is equivalent to solving a nonlinear integral equation; this simplification is what makes the problem tractable. Introducing a function $Y(Q)$ by $F(Q)=Y^{m+1}(Q)$, with $m+1$ the coordination number of the tree, I write the nonlinear integral equation in the form

$$Y(Q)=\int d\mu(Q')L(Q;Q')D(Q')Y^m(Q')\,, \tag{5.1}$$

where the kernel L, the symmetry-breaking term D, and the invariant measure $d\mu(Q)$ were defined in Ref. 15. Due to the graded symmetry of L and D, the solution Y is a function only of the "eigenvalues" λ_1 and λ_2.[10,15]

Eq. (5.1) can be reduced to an integral equation in just these two variables by applying the theorem (4.3) to the Q' integral. Using that D and Y depend only on λ_1' and λ_2', we easily find

$$Y(\lambda_1,\lambda_2)=\exp[-\gamma(\lambda_1-\lambda_2)]+\int_1^\infty d\lambda_1'\int_{-1}^{+1}d\lambda_2'(\lambda_1'-\lambda_2')^{-2}L_4(\lambda_1,\lambda_2;\lambda_1',\lambda_2')D(\lambda_1',\lambda_2')Y^m(\lambda_1',\lambda_2')\,, \tag{5.2}$$

where L_4 is obtained by expanding $L(Q;Q')$ and integrating over ϕ_1,ϕ_2; see Sec. IV and also Ref. 15:

$$L_4(\lambda_1,\lambda_2;\lambda_1',\lambda_2')=e^{-\gamma(\lambda_1\lambda_1'-\lambda_2\lambda_2')}\left\{\frac{\gamma}{4}[\,|\mu_1\mu_1'|I_1(\gamma|\mu_1\mu_1'|)I_0(\gamma|\mu_2\mu_2'|)+|\mu_2\mu_2'|I_0(\gamma|\mu_1\mu_1'|)I_1(\gamma|\mu_2\mu_2'|)]\right.$$
$$+\frac{\gamma^2}{4}\{[(\lambda_1-\lambda_2)^2(\lambda_1'-\lambda_2')^2+|\mu_1\mu_2'|^2+|\mu_2\mu_1'|^2$$
$$+\tfrac{1}{2}|\mu_1\mu_1'|^2+\tfrac{1}{2}|\mu_2\mu_2'|^2]I_0(\gamma|\mu_1\mu_1'|)I_0(\gamma|\mu_2\mu_2'|)-2(\lambda_1-\lambda_2)(\lambda_1'-\lambda_2')$$
$$\times[\,|\mu_1\mu_1'|I_1(\gamma|\mu_1\mu_1'|)I_0(\gamma|\mu_2\mu_2'|)-|\mu_2\mu_2'|I_0(\gamma|\mu_1\mu_1'|)I_1(\gamma|\mu_2\mu_2'|)]$$
$$\left.+\tfrac{1}{2}|\mu_1\mu_1'|^2I_2(\gamma|\mu_1\mu_1'|)I_0(\gamma|\mu_2\mu_2'|)+\tfrac{1}{2}|\mu_2\mu_2'|^2I_0(\gamma|\mu_1\mu_1'|)I_2(\gamma|\mu_2\mu_2'|)\}\right]. \tag{5.3}$$

Efetov[10] has given an integral equation for Y that is equivalent to Eqs. (5.2) and (5.3) but is much more compact. The equivalence between the two forms of writing the equation can be shown by using a standard recurrence relation among modified Bessel functions, $I_2(z)=I_0(z)-(2/z)I_1(z)$.

Efetov[10] has analyzed Eq. (5.2) for $\gamma\to 0$ and $\gamma\to\infty$, and he found that the solutions Y in these limits correspond to localized and extended states, respectively. It then follows that the system undergoes a phase transition ("mobility transition") at some finite value of γ, $\gamma=\gamma_c$. In the sequel, I will reanalyze the critical behavior of all two-point Green's functions at this point.

A. Localized states revisited

It is suggested by the general discussion of Sec. IV (and confirmed by the explicit calculations of Refs. 10 and 15) that $\lim_{\epsilon\to 0}Y(\lambda_1,\lambda_2;\gamma,\epsilon)=1$ for all λ_1 and λ_2 when states are localized ($\gamma<\gamma_c$). For ϵ infinitesimal, the solution Y can be written as $Y(\lambda_1,\lambda_2)=\psi(\epsilon\lambda_1)$. Y depends only on the combination $\epsilon\lambda_1$ because the symmetry-breaking term has the form $D=\exp[-2\epsilon(\lambda_1-\lambda_2)]$. By making the substitution $\lambda_1=\epsilon^{-1}e^t$, letting $\epsilon\to 0$, and performing the integral over λ_2, we can reduce Eq. (5.2) to an integral equation in the variable t only:

$$\psi(t)=\int_{-\infty}^{+\infty}dt'L_\gamma(t-t')d(t')\psi^m(t')\,, \tag{5.4}$$

where $d(t)=\exp(-2e^t)$, $L_\gamma=\exp[\frac{1}{2}(t-t')]l_\gamma(t-t')$ and l_γ was defined in Ref. 15. The critical behavior of the solution ψ of Eq. (5.4) near $\gamma=\gamma_c$ was studied in Refs. 10 and 15. It was found that the solution tends toward a definite limiting shape as $\gamma\to\gamma_c$ and then abruptly collapses to the trivial solution $\psi=0$ at $\gamma=\gamma_c$. This implies in particular that $\Gamma^{\rm loc}(0;\gamma)$ goes discontinuously to zero at the critical point.

In Ref. 15 also $\Gamma^{\rm loc}(x;\gamma)$ for $x\neq 0$ was considered, and an exponential decay with $|x|$ was found. Unfortunately, I failed to interpret these numerical results correctly, not realizing that $\Gamma^{\rm loc}(x;\gamma)$ (and, in fact, all correlation functions) cannot but decrease exponentially, due to the exponential growth with distance of the number of neighbors on the Bethe lattice. The objective of this subsection is to show that the "weighted" density-density correlation function (weighted by the number of neighbors) does exhibit interesting behavior.

For the present purpose, it is convenient to organize the evaluation of $\Gamma^{\rm loc}$ in the following way:

$$\delta\psi^{(0)}(t)=e^t\psi(t)\,,$$
$$\delta\psi^{(1)}(t)=(m+1)\int_{-\infty}^{+\infty}dt'L_\gamma(t-t')\times d(t')\psi^{m-1}(t')\delta\psi^{(0)}(t')\,,$$
$$\delta\psi^{(2)}(t)=m\int_{-\infty}^{+\infty}dt'L_\gamma(t-t')d(t')\psi^{m-1}(t')\delta\psi^{(1)}(t')\,,$$
$$\vdots \tag{5.5}$$
$$\delta\psi^{(p)}(t)=m\int_{-\infty}^{+\infty}dt'L_\gamma(t-t')d(t')\psi^{m-1}(t')\delta\psi^{(p-1)}(t')\,,$$
$$N(|x-y|)\Gamma^{\rm loc}(x-y;\gamma)=2\int_{-\infty}^{+\infty}dt\,d(t)\psi^m(t)\delta\psi^{(p)}(t)\,.$$

Here, $p=|x-y|$ is the distance between the sites x and y, and $N(p)$,

$$N(0)=1,\ N(1)=m+1,\ N(2)=(m+1)m,\ldots$$
$$N(p)=(m+1)m^{p-1},\ldots \tag{5.6}$$

denotes the number of neighbors at a distance p.

There exists a remarkable connection between the above sequence of equations and the integral equation (5.4). Suppose that we were to perturb the solution of Eq. (5.4) by an infinitesimal amount, $\psi\to\psi+\delta\psi^{(0)}$. The evolution of this perturbation under iteration of Eq. (5.4) is governed by Eqs. (5.5), for the linear operator $M(t;t')=mL_\gamma(t-t')Z(t')$, $Z(t')=d(t')\psi^{m-1}(t')$, entering these equations is precisely the operator obtained by linearizing Eq. (5.4) around its solution ψ. There are two immediate conclusions that can be drawn from this observation. On the one hand, the required stability (for $\gamma<\gamma_c$) of the solution ψ with respect to small perturbations implies that $\delta\psi^{(p)}(t)$ goes to zero for all t as $p\to\infty$. On the other hand, the observed increase[10,15] in the number of steps needed to reach convergence in the iterative solution of Eq. (5.4) is a direct signal of the appearance of long-range correlations in $\Gamma^{\rm loc}(x-y;\gamma)$.

The evolution of the "perturbation" $\delta\psi^{(0)}(t)=e^t\psi(t)$ is best analyzed by introducing the eigenfunctions of M. It is rather difficult to get a handle on these eigenfunctions

by analytical means but, fortunately, this will prove unnecessary. Notice, first of all, that the operator M has a continuous spectrum since the integration domain is unbounded from below and

$$M(t;t') \to mL_\gamma(t-t') = me^{(t-t')/2} l_\gamma(t-t') \tag{5.7}$$

for $t \to -\infty$. Notice also that M is not invariant under the interchange of its arguments t and t'. This means that M is not Hermitian in an appropriate functional space, and its eigenfunctions are not orthogonal with respect to the unit weight in this space. However, if $\psi_\rho(t)$ is an eigenfunction of M, then $e^{-t/2}\sqrt{Z(t)}\psi_\rho(t)$ is an eigenfunction of the *symmetric* kernel $\sqrt{Z(t)}l_\gamma(t-t')\sqrt{Z(t')}$. We therefore expect the $\psi_\rho(t)$ to satisfy orthogonality relations like

$$\int_{-\infty}^{+\infty} dt\, e^{-t} Z(t) \psi_\rho(t) \psi_{\rho'}(t) = \frac{\delta(\rho-\rho')}{W_\rho} , \tag{5.8}$$

where W_ρ depends on how we normalize the ψ_ρ. The corresponding completeness relation in the appropriate functional space takes the form

$$e^{-t} Z(t) \int_0^\infty W_\rho d\rho\, \psi_\rho(t) \psi_\rho(t') = \delta(t-t') . \tag{5.9}$$

Due to Eq. (5.7), the asymptotic behavior of the eigenfunctions for $t \ll 0$ is given by

$$\psi_\rho(t) \simeq e^{t/2} \sin[-\rho t + \delta(\rho)] . \tag{5.10}$$

With the normalization chosen in (5.10), W_ρ will be regular and nonvanishing at $\rho = 0$ (i.e., $\lim_{\rho\to 0} W_\rho = W_0 > 0$), as is suggested by the orthonormality of the functions $\sqrt{2/\pi} e^{t/2} \sin(\rho t)$ over the interval $(-\infty, 0]$ with weight e^{-t}. We also need to know how the "phase shift" $\delta(\rho)$ behaves at $\rho = 0$. The continuity of the eigenfunctions in conjunction with the fact that the term $Z(t)$ forces them to approach zero near $t = 0$, requires that the ratio $\delta(\rho)/\rho$ tend toward a definite limit as $\rho \to 0$. In other words,

$$\delta(\rho) \simeq \rho t_0 + O(\rho^2) . \tag{5.11}$$

These considerations put us in a position where we can evaluate the long-range part of Γ^{loc}. To do so, we first "Fourier analyze" the initial perturbation $\delta\psi^{(0)}(t)$ in terms of the eigenfunctions of M:

$$\delta\psi_\rho^{(0)} = \int_{-\infty}^{+\infty} dt\, e^{-t} Z(t) \psi_\rho(t) \delta\psi^{(0)}(t) . \tag{5.12}$$

The evolution of $\delta\psi_\rho^{(0)}$ is now obtained simply by successive multiplication with the corresponding eigenvalue $m\kappa_\rho$ of M: $\delta\psi_\rho^{(p)} = (m+1) m^{p-1} \kappa_\rho^p \delta\psi_\rho^{(0)}$. We finally invert (5.12) using the completeness relation (5.9) and insert the result into the last of Eqs. (5.5). This gives

$$N(|x|)\Gamma^{\text{loc}}(x;\gamma) = \frac{m+1}{m} \int_0^\infty W_\rho d\rho (m\kappa_\rho)^{|x|} \delta\psi_\rho^{(0)} \times \left[2 \int_{-\infty}^{+\infty} dt\, d(t) \psi^m(t) \psi_\rho(t) \right] . \tag{5.13}$$

The discussion now focuses on the expression $(m\kappa_\rho)^{|x|}$ for $|x| \gg 1$. The eigenvalue κ_ρ can be obtained by acting on $\psi_\rho(t)$ with $L_\gamma(t-t')$ in the asymptotic domain $t \ll 0$. Using Eqs. (5.7), (5.10), and the fact that $l_\gamma(t-t')$ is even in $t-t'$ we find that

$$m\kappa_\rho = \exp[-a_0 - a_2\rho^2 + O(\rho^4)] . \tag{5.14}$$

Note that κ_ρ is related to the eigenvalue $\omega_\gamma(\theta)$ used in Ref. 15 by $\kappa_\rho = \omega_\gamma(\frac{1}{2} \pm i\rho)$. The equation for the critical point γ_c derived in Refs. 10 and 15 takes the form $m\kappa_0 = 1$, which implies that $a_0 = \bar{a}_0(\gamma - \gamma_c) + O((\gamma-\gamma_c)^2)$ near $\gamma = \gamma_c$. According to the central limit theorem of statistics, we may now put

$$(m\kappa_\rho)^{|x|} \simeq e^{-a_0|x|} e^{-a_2|x|\rho^2} \quad (|x| \gg 1) . \tag{5.15}$$

Finally, we use that $\delta\psi_\rho^{(0)}$ and the integral in large parentheses in Eq. (5.13) are both linear in ρ for ρ small. [Both statements derive from the asymptotic form (5.10) and the property (5.11) of $\delta(\rho)$.] The resulting expression for the (weighted) density-density correlation function is

$$N(|x|)\Gamma^{\text{loc}}(x;\gamma) \simeq \text{const} \times e^{-a_0|x|} \int_0^\infty d\rho\, \rho^2 e^{-a_2|x|\rho^2} = C|x|^{-3/2} \exp(-|x|/\zeta) , \tag{5.16}$$

where $\zeta^{-1} = a_0 \propto |\gamma - \gamma_c|$ vanishes linearly at the mobility edge.[19]

It should be pointed out that the power law proportional to $|x|^{-\delta}$, $\delta = \frac{3}{2}$, is consistent with $\Gamma^{\text{loc}}(0;\gamma_c -) > 0$ and the sum rule (4.12),

$$1 = \sum_x \Gamma^{\text{loc}}(x;\gamma) = \sum_{p=0}^\infty N(p) \Gamma^{\text{loc}}(p;\gamma) . \tag{5.17}$$

For $\Gamma^{\text{loc}}(0;\gamma_c -) > 0$ the sum rule can only be conserved if $\delta > 1$. Conversely, a smaller value of the exponent, $0 < \delta \leq 1$, would necessitate the continuous vanishing of $\Gamma^{\text{loc}}(0;\gamma)$ at $\gamma = \gamma_c$. It should be added that the scaling behavior (5.16) is completely confirmed by the numerical results given in Ref. 15 and their subsequent extension.

B. Extended states

After this addendum to previous work, we turn our attention to the region of extended states. Analysis of this phase will require considerably more effort, as is suggested by a quick glance at Eqs. (5.2) and (5.3), or at Eqs. (17) and (18) of Ref. 10. It will prove useful to develop our intuition about the properties of this integral equation by looking at numerical results first.

According to a discussion in Sec. IV, we expect the global U(1,1/2) symmetry of the system to remain broken as $\epsilon \to 0$ for $\gamma > \gamma_c$. The expected finiteness of all two-point functions in this limit leads us to seek solutions Y that decay as $\lambda_1 \to \infty$. If such decaying solutions exist (and it will be demonstrated below that this is indeed the case), then we can take the limit $\epsilon \to 0$ right under the integral sign in Eq. (5.2), for, under these conditions, the convergence of $Y(\lambda_1,\lambda_2;\gamma,\epsilon)$ to $Y(\lambda_1,\lambda_2;\gamma,\epsilon=0)$ is *uniform* in λ_1 and λ_2. We accordingly put $D \to 1$ in Eq. (5.2).

It should be pointed out here that the resulting equation has $Y \equiv 1$ as a solution for *all* values of γ. This follows from the fact that the kernel $L(Q;Q')$ is normalized to

unity, $\int d\mu(Q')L(Q;Q')=1$.[15] However, for $\gamma>\gamma_c$ (extended states) this constant solution is *unstable* with respect to the addition of a small symmetry-breaking term, i.e., it cannot be reached by the procedure of first taking $\epsilon>0$, then solving the integral equation (5.2), and finally letting $\epsilon\to 0$. The constant solution must therefore be discarded for $\gamma>\gamma_c$.

The nonlinear integral equation for Y, Eq. (5.2), is the appropriate equation to use when λ_1 is large. However, as γ approaches γ_c from above, the solution Y will be seen to approximate 1 very closely for a large range of values of λ_1, $1\le\lambda_1\lesssim\Lambda$. In this range it is therefore better to consider an equivalent integral equation for the deviation of Y from unity. This greatly improves the numerical stability, and it also helps the analytical discussion in Sec. V B 2. Following Efetov,[10] we put $Y=1-u(\lambda_1-\lambda_2)$. We also use the relation quoted in the text below Eq. (5.3), and the identity

$$\exp[-\gamma(\lambda_1-\lambda_2)]+\int_1^\infty d\lambda_1'\int_{-1}^{+1}d\lambda_2'(\lambda_1'-\lambda_2')^{-2}L_4(\lambda_1,\lambda_2;\lambda_1',\lambda_2')=1\ , \tag{5.18}$$

which derives from the normalization property of the kernel $L(Q;Q')$. Equation (5.2) with $\epsilon=0$ can then be rewritten as a nonlinear integral equation for the quantity u:

$$u=\hat{K}*[mu-X(u)]\ . \tag{5.19}$$

Here, $\hat{K}$ is an integral operator with kernel

$$(\hat{K}*u)(\lambda_1,\lambda_2)=\int_1^\infty d\lambda_1'\int_{-1}^{+1}d\lambda_2'K(\lambda_1,\lambda_2;\lambda_1',\lambda_2')u(\lambda_1',\lambda_2')\ , \tag{5.20a}$$

$$K(\lambda_1,\lambda_2;\lambda_1',\lambda_2')=\frac{\gamma^2}{2}e^{-\gamma(\lambda_1\lambda_1'-\lambda_2\lambda_2')}[(\lambda_1\lambda_1'+\lambda_2\lambda_2')I_0(\gamma\,|\,\mu_1\mu_1'\,|\,)I_0(\gamma\,|\,\mu_2\mu_2'\,|\,)$$
$$-\,|\,\mu_1\mu_1'\,|\,I_1(\gamma\,|\,\mu_1\mu_1'\,|\,)I_0(\gamma\,|\,\mu_2\mu_2'\,|\,)+|\,\mu_2\mu_2'\,|\,I_0(\gamma\,|\,\mu_1\mu_1'\,|\,)I_1(\gamma\,|\,\mu_1\mu_1'\,|\,)]\ , \tag{5.20b}$$

and X is a polynomial in u and the difference $(\lambda_1-\lambda_2)$,

$$X(u)=\{[1-u(\lambda_1-\lambda_2)]^m+mu(\lambda_1-\lambda_2)-1\}/(\lambda_1-\lambda_2)$$
$$=\sum_{p=2}^{m}c_p(m)u^p(\lambda_1-\lambda_2)^{p-1},\quad c_p(m)=(-1)^p\begin{bmatrix}m\\p\end{bmatrix}\ . \tag{5.21}$$

Some properties of X have been discussed by Efetov.[10] (Note that X is denoted by F in Ref. 10.) The kernels L_4 and K are connected by the relation

$$L_4(\lambda_1,\lambda_2;\lambda_1',\lambda_2')=K(\lambda_1,\lambda_2;\lambda_1',\lambda_2')(\lambda_1-\lambda_2)(\lambda_1'-\lambda_2')\ . \tag{5.22}$$

K has the interesting property[10] that it can be expressed solely in terms of the invariant product $\hat{\lambda}_1\cdot\hat{\lambda}_1'=\lambda_1\lambda_1'-\frac{1}{2}[\mu_1^*\mu_1'+(\mu_1^*)'\mu_1]$ on the hyperboloid $SU(1,1)/U(1)\simeq H_2$, and the invariant product $\hat{\lambda}_2\cdot\hat{\lambda}_2'=\lambda_2\lambda_2'+\frac{1}{2}[\mu_2^*\mu_2'+(\mu_2^*)'\mu_2]$ on the sphere $SU(2)/U(1)\simeq S_2$:

$$K(\lambda_1,\lambda_2;\lambda_1',\lambda_2')=\frac{\gamma^2}{2}\int_0^{2\pi}\frac{d\phi_1'}{2\pi}\int_0^{2\pi}\frac{d\phi_2'}{2\pi}e^{-\gamma(\hat{\lambda}_1\cdot\hat{\lambda}_1'-\hat{\lambda}_2\cdot\hat{\lambda}_2')}$$
$$\times(\hat{\lambda}_1\cdot\hat{\lambda}_1'+\hat{\lambda}_2\cdot\hat{\lambda}_2')\ . \tag{5.23}$$

An important consequence of this invariance is that

$$\int_1^\infty d\lambda_1'\int_{-1}^{+1}d\lambda_2'K(\lambda_1,\lambda_2;\lambda_1',\lambda_2')$$
$$=\int_1^\infty d\lambda_1'\int_{-1}^{+1}d\lambda_2'K(1,1;\lambda_1',\lambda_2')$$
$$=\frac{\gamma^2}{2}\int_1^\infty d\lambda_1\int_{-1}^{+1}d\lambda_2e^{-\gamma(\lambda_1-\lambda_2)}(\lambda_1+\lambda_2)=1\ . \tag{5.24}$$

From the invariance of K we must not conclude, however, that the solution too is invariant on $H_2\otimes S_2$, for the symmetry on this space is explicitly broken by the factors $(\lambda_1-\lambda_2)$ appearing in the definition of X. We deduce from the properties of Y,[10,15] $Y(1,1)=1, 1\ge Y>0$, and $\lim_{\lambda_1\to\infty}Y(\lambda_1,\lambda_2)=0$, that $u(1,1)=c<\infty$, $u<(\lambda_1-\lambda_2)^{-1}\forall\lambda_1,\lambda_2$, and $u\to\lambda_1^{-1}$ as $\lambda_1\to\infty$.

1. Numerical solution

The numerical solution of Eq. (5.2) or, equivalently, of Eq. (5.19) is a nontrivial task which is, however, somewhat simplified by the positive definiteness of the kernels L_4 and K. Equations (5.2) and (5.19) are integral equations in *two* variables, one of which is noncompact, and this places considerable demands on both storage and computation time. As the properties of these integral equations became unravelled, I gradually settled for the numerical method outlined below.

For reasons indicated earlier, I chose to split up the interval $1\le\lambda_1<\infty$ into two regions. In the first region, $1\le\lambda_1<\Lambda$, Eq. (5.19) was used, and in the second region,

$\Lambda \leq \lambda_1 < \infty$, Eq. (5.2) with Eq. (5.3) or, rather, Efetov's simple form[10] of this equation. The point Λ was chosen so as to satisfy the condition $Y(\Lambda,\lambda_2) \simeq 0.5$. [We will see that as $\gamma \to \gamma_c$, Λ becomes large and $Y(\Lambda,\lambda_2)$ independent of λ_2.] Both equations (5.19) and (5.2) were discretized, with mesh points chosen to be equally spaced in λ_2, and in the variable $t = \ln\lambda_1$. Typically, 5 to 10 points were used in λ_2, and several hundred in t. Of course, only a finite number is needed since we are looking for solutions that decay as $\lambda_1 \to \infty$. The adequacy of the discretization procedure was checked by verifying the relations (5.18) and (5.24). Actually, the first of these relations is not easy to preserve under discretization, due to an (integrable) singularity at $\lambda_1' = \lambda_2' = 1$. This does not cause any numerical problems, however, as it is Eq. (5.19) and not Eq. (5.2) that is used for small values of λ_1. (With increasing λ_1, contributions from the singular point become weaker because of the finite range of L_4.) The integral equations were solved iteratively, with an initial function $Y_0(\lambda_1,\lambda_2) = \exp[-\beta(\lambda_1 - \lambda_2)]$. This initial function is suggested by Efetov's analysis[10] of the metallic limit $\gamma \to \infty$.

Let me now describe the features of the solution u to Eq. (5.19) as found by this numerical procedure; all of these features will be explained qualitatively, and to a large extent also quantitatively, by the analysis given below. To begin, nontrivial (i.e., nonvanishing) solutions to Eq. (5.19) do indeed exist for $\gamma > \gamma_c$ [i.e., technically speaking, the global U(1,1/2) symmetry is indeed spontaneously broken]. These solutions decrease monotonically with increasing λ_1. For a stretch of λ_1 values excluding the vicinity of the origin and a small region near the crossover point Λ, the solution u behaves roughly as $\sim \lambda_1^{-1/2}$. Beyond the crossover point, u rapidly attains the asymptote $\sim \lambda_1^{-1}$. As the critical point is approached, $u(\lambda_1,\lambda_2)$ is pushed down (for $\lambda_1 \lesssim \Lambda$ only) and, consequently, Λ moves to larger values of λ_1. The region in which the crossover from the behavior $\sim \lambda_1^{-1/2}$ to the asymptotic behavior $\sim \lambda_1^{-1}$ occurs remains of finite size as $\gamma \to \gamma_c$. These features are evident from Fig. 1 which shows $\ln u(\lambda_1,1)$ as a function of $t = \ln\lambda_1$ for three different values of γ. Using Eqs. (3.2b) and (4.4b) to calculate the diagonal part of Γ^{ext}, I find a violent singularity in this observable; for example, for $m = 2$ and $10^2 \times \gamma = 7.10,\ 7.05,\ 7.00,\ 6.95,\ 6.90$, I obtain $\Gamma^{\text{ext}}(0;\gamma) = 2.06 \times 10^{22},\ 2.52 \times 10^{24},\ 1.69 \times 10^{27},\ 2.44 \times 10^{31},\ 2.64 \times 10^{38}$. (The critical point lies in this case at $\gamma_c = 6.803 \times 10^{-2}$.)

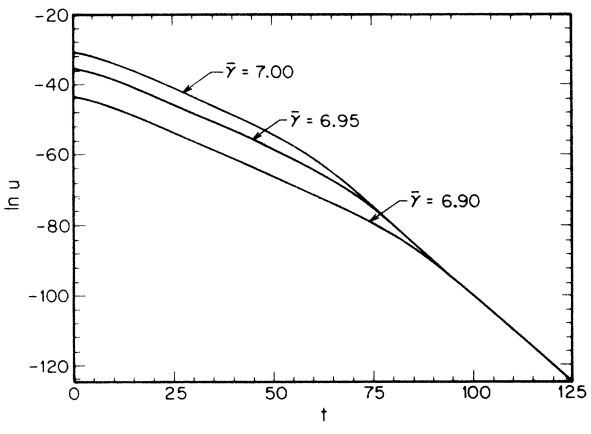

FIG. 1. The logarithm of the solution $u(\lambda_1,1)$ of Eq. (5.19) as a function of $t = \ln\lambda_1$ for $m = 2$ and three different values of $\bar{\gamma} \equiv 10^2 \times \gamma$. The critical point lies at $\gamma_c = 6.803 \times 10^{-2}$.

These numerical results are in striking contrast to the analysis of Efetov[10] who claims that such long-ranged solutions as are shown in Fig. 1 cannot exist, with the peculiar consequence that Γ^{ext} remains finite as $\gamma \to \gamma_c$. Therefore, several tests were performed in order to check the numerical stability of the solution with respect to changes in the initial function and the discretization procedure used. The outcome of all these tests was satisfactory, suggesting that the appearance of long-ranged solutions is in fact a genuine property of the integral equation and not a spurious effect caused by discretization. In order for the present results to be completely convincing, I should, of course, point out what went wrong in Efetov's analysis. This will be done in Sec. V B 2.

I also found that close to the critical point the solution u becomes independent of the variable λ_2 for *all* values of λ_1. This is easy to understand. Recall that the S_2 symmetry of the integral equation (5.19) is violated only by the difference $(\lambda_1 - \lambda_2)$ hidden in the nonlinear term X. For $\lambda_1 \gg 1$, this difference can always be replaced by λ_1; this explains why the solution u at *large* λ_1 is *always* independent of λ_2, no matter whether the system is close to the critical point or not. In the vicinity of $\gamma = \gamma_c$, however, an additional feature arises: $u(1,1)$ approaches zero and, consequently, the nonlinear term $X = c_2 u^2(\lambda_1 - \lambda_2) + O(u^3(\lambda_1 - \lambda_2)^2)$ becomes negligible compared to the term mu as long as $u(\lambda_1 - \lambda_2) \ll 1$. As $\gamma \to \gamma_c$, the integral equation thus becomes effectively invariant on S_2 for *all* values of λ_1. Using the arguments given in Appendix B of Ref. 15, one then concludes that the effective S_2 symmetry of the integral equation carries over to the solution u.

It is now clear that the critical behavior at the phase transition point is governed exclusively by the noncompact variable, λ_1. The only role of the compact variable, λ_2, is to ensure that certain integrals [see Eqs. (5.18) and (5.24)] are properly normalized to unity. This should be contrasted with a recent theory of the quantized Hall effect,[5,6] which proposes to explain the existence of extended states in two-dimensional disordered systems by topological arguments pertaining to the compact sector, i.e., the variable λ_2. Clearly, the transition mechanism assumed in Refs. 5 and 6 is completely unrelated to the one found here.

Due to the effective S_2 symmetry near $\gamma = \gamma_c$, the integral equation (5.19) can be reduced to an equation in a *single* variable by averaging the kernel K over S_2 and putting $u(\lambda_1,\lambda_2) \to u(\lambda_1)$. This greatly accelerates the numerical computation. It also allows a finer mesh size to be used and makes it possible to go much closer to the critical point.

2. Analytical discussion

It has been argued above, that Eq. (5.19) effectively reduces to an integral equation on the hyperboloid $H_2 \simeq SU(1,1)/U(1)$ near the critical point $\gamma=\gamma_c$. To begin the analytical discussion of the present subsection, I write the reduced equation as

$$u(Q)=\int d\mu(Q')K(Q;Q')[mu(Q')-X(Q',u)] , \quad (5.25a)$$

$$K(Q;Q')=e^{\frac{1}{2}\gamma \mathrm{tr} QQ'}[c_0(\gamma)-c_1(\gamma)\mathrm{tr}QQ'] , \quad (5.25b)$$

$$c_0(\gamma)=\gamma\cosh\gamma-\sinh\gamma, \quad c_1(\gamma)=\tfrac{1}{2}\gamma\sinh\gamma . \quad (5.25c)$$

Here, Q is *not* the graded matrix defined in Eq. (2.1) but is instead given by

$$Q=\begin{bmatrix} -i\lambda & \mu \\ \mu^* & +i\lambda \end{bmatrix}, \quad \lambda^2-|\mu|^2=1, \quad \mu=|\mu|e^{i\phi} ;$$

$$d\mu(Q)=d\lambda\, d\phi/2\pi . \quad (5.26)$$

[For notational reasons, I have written $u=u(Q)$ although u actually depends only on λ_1, and I have put $\lambda_1\to\lambda,\mu_1\to\mu$.] Formally,

$$Q=-iT\sigma_3T^{-1}, \quad T\in SU(1,1), \quad Q\in SU(1,1)/U(1) . \quad (5.27)$$

The present way of writing the integral equation emphasizes the H_2 symmetry of the problem and is convenient for subsequent analysis.

In the region $\lambda \ll \Lambda$ (recall that $\Lambda\to\infty$ as $\gamma\to\gamma_c$), where the nonlinear term X is small compared to mu, the solution u is forced into an eigenfunction of K with eigenvalue $\kappa\to m^{-1}$. It is therefore useful to introduce[10] a (complete) set of eigenfunctions of the kernel K. These can be obtained from the irreducible representations of the group SU(1,1), which is seen as follows. Suppose that we are given a set of functions $\mathscr{P}_\nu(Q)$ that satisfy an addition theorem of the form

$$\mathscr{P}_\nu(TQ'T^{-1})=\sum_n \mathscr{P}_{\nu n}(Q)\mathscr{P}^*_{\nu n}(Q') , \quad (5.28a)$$

$$Q=-iT\sigma_3T^{-1}, \quad Q'=-iT'\sigma_3(T')^{-1} , \quad (5.28b)$$

$$\mathscr{P}_{\nu n}(Q)=e^{in\phi}C_{\nu n}(\lambda), \quad \mathscr{P}_{\nu 0}(Q)=\mathscr{P}_\nu(Q) . \quad (5.28c)$$

Because K is invariant on $H_2\simeq SU(1,1)/U(1)$, functions with this property are eigenfunctions of K:

$$\int d\mu(Q'')K(Q;Q'')\mathscr{P}_\nu(Q'') = \int d\mu(Q')K(-i\sigma_3;Q')\mathscr{P}_\nu(TQ'T^{-1})=k_\nu\mathscr{P}_\nu(Q) , \quad (5.29)$$

where

$$k_\nu=\int d\mu(Q')K(-i\sigma_3;Q')\mathscr{P}^*_{\nu 0}(Q') . \quad (5.30)$$

Now the addition theorem (5.28a) is precisely the representation property for group multiplication in SU(1,1),

$$D^\nu_{mm'}(TT')=\sum_n D^\nu_{mn}(T)D^\nu_{nm'}(T') ,$$

with $m=m'=0$. [Note that the matrix $Q''\in SU(1,1)/U(1)$ which corresponds to the product TT' is $Q''=-iTT'\sigma_3(TT')^{-1}=TQ'T^{-1}$.] We thus conclude that the functions $\mathscr{P}_\nu(Q)=D^\nu_{00}(T)$ are the desired eigenfunctions of K.

Following Vilenkin,[20] I write the 00 matrix elements of the irreducible representations of SU(1,1) in the integral form

$$D^\nu_{00}(T)=\mathscr{P}_\nu(Q)=\mathscr{P}_\nu(\lambda) = \int_0^{2\pi}\frac{d\theta}{2\pi}(\cosh\tau+\sinh\tau\cos\theta)^\nu, \quad \lambda=\cosh\tau . \quad (5.31)$$

Among the functions $\mathscr{P}_\nu$ (also known as Legendre functions), those with index $\nu=-\frac{1}{2}+i\rho$, ρ real (the so-called cone functions), play a special role: they correspond to *unitary* representations of SU(1,1), they decay as $\lambda\to\infty$, and they are orthogonal and complete in the space of square-integrable functions $F(\lambda)$ over the interval $[1,\infty)$. The orthogonality and completeness relations for the $\mathscr{P}_{-1/2+i\rho}(\lambda)$ can be found in Ref. 21. Here, we need only their asymptotic form for $\lambda\gg1$ and $\rho\ll1$, which is derived in the Appendix:

$$\mathscr{P}_{-1/2+i\rho}(\cosh\tau)\simeq\frac{2}{\pi\rho}e^{-\tau/2}\sin(\rho\tau) . \quad (5.32)$$

Note also that the $\mathscr{P}_{-1/2+i\rho}(\lambda)$ are real and even functions of ρ, $\mathscr{P}^*_{-1/2+i\rho}(\lambda)=\mathscr{P}_{-1/2+i\rho}(\lambda)=\mathscr{P}_{-1/2-i\rho}(\lambda)$. Using Eq. (5.29), the integral equation for u can now be transformed into

$$(m-\kappa_\rho^{-1})u_\rho=X_\rho , \quad (5.33)$$

where

$$u_\rho=\int_1^\infty d\lambda\,\mathscr{P}_{-1/2+i\rho}(\lambda)u(\lambda) , \quad X_\rho=\int_1^\infty d\lambda\,\mathscr{P}_{-1/2+i\rho}(\lambda)X(\lambda,u(\lambda)) , \quad (5.34)$$

and $\kappa_\rho=k_{-1/2+i\rho}$.

Due to the positivity of K, u and X, and the oscillatory nature of $\mathscr{P}_{-1/2+i\rho}(\lambda)$ [see, e.g., Eq. (5.32)], the functions κ_ρ, u_ρ, and X_ρ have maxima at $\rho=0$. [Because $\mathscr{P}_{-1/2+i\rho}(\lambda)$ is analytic in the complex ρ plane, the point $\rho=0$ is in fact a *saddle point* of these functions.] The behavior of u_ρ can be readily obtained from the numerical results for $u(\lambda)$ shown in Fig. 1. One finds that u_ρ is an oscillating function with an envelope that decreases as a power as $\rho\to\infty$. A crude estimate of the smallest zero of u_ρ, $\rho=\rho_1$, can be had by approximating $u(\lambda)$ by $u(\lambda)\simeq u(1)\mathscr{P}_{-1/2}(\lambda)$ for $\lambda\le\Lambda$, and by $u(\lambda)\simeq0$ for $\lambda>\Lambda$. Using also the asymptotic form (5.32), we then get with $\Lambda\equiv\cosh\tau_c$:

$$u_\rho=u(1)\int_1^\Lambda d\lambda\,\mathscr{P}_{-1/2}(\lambda)\mathscr{P}_{-1/2+i\rho}(\lambda) \simeq\frac{2u(1)}{\pi^2\rho^3}\int_0^{\rho\tau_c}dx\,x\sin x , \quad (5.35)$$

from which we infer that ρ_1 is located in the interval $\pi/\tau_c<\rho_1<2\pi/\tau_c$. [I mention in passing that due to the asymptotic limit $u(\lambda)\simeq\lambda^{-1}$ for $\lambda\to\infty$, the function u_ρ is

analytic inside the strip $-\frac{1}{2}<\mathrm{Im}\rho<\frac{1}{2}$ and has poles at $\rho=\pm\frac{1}{2}i$. The existence of these poles is of no consequence for the present analysis and a corresponding term, originating from the integration range $\tau>\tau_c$, was omitted from Eq. (5.35).]

An *additional* node on the left-hand side (lhs) of the transformed integral equation is induced by the presence of the factor $(m-\kappa_\rho^{-1})$; this has important consequences as will be discussed next. As was stated earlier, κ_ρ takes its maximum value for real ρ at $\rho=0$. The decrease of κ_ρ as $\rho\to\infty$ is qualitatively slower than that of u_ρ because $K(-i\sigma_3;Q)$ is a "short-ranged" function (i.e., K falls off exponentially with λ), while $u(\lambda)$ becomes "long ranged" at the critical point [i.e., $u(\lambda)\propto\lambda^{-1/2}$ for $\lambda\ll\Lambda$]. In the region of extended states, the maximum eigenvalue κ_0 exceeds m^{-1} and $\kappa_0\to m^{-1}$ as $\gamma\to\gamma_c$.[10] For ρ small and γ close to γ_c, $m-\kappa_\rho^{-1}$ is thus expressed as

$$m-\kappa_\rho^{-1}=b_0-b_2\rho^2+O(\rho^4)\ , \tag{5.36a}$$

$$b_0=\bar{b}_0(\gamma-\gamma_c)+O((\gamma-\gamma_c)^2)\ , \tag{5.36b}$$

where $\bar{b}_0$ and b_2 are noncritical positive constants that depend on the coordination number of the tree. (Note that b_0 and b_2 are connected with the quantities a_0 and a_2 in Sec. V A through $b_0=-ma_0$, $b_2=+ma_2$.) The function $m-\kappa_\rho^{-1}$ has a zero at $\rho_0\simeq(b_0/b_2)^{1/2}$. The critical behavior of b_0 forces this zero to approach the origin in ρ as $\gamma\to\gamma_c$; we shall see that this is the mechanism that drives the phase transition from extended to localized states.

If Eq. (5.33) is to be satisfied, then the node at $\rho=\rho_0$ induced by the factor $(m-\kappa_\rho^{-1})$ on the lhs must be matched by a corresponding node on the rhs. To see what effect the location of this node has on the solution u, we write X_ρ in the form

$$X_\rho\simeq\frac{1}{\pi\rho}\int_0^\infty d\tau(e^{\tau/2}X)\sin(\rho\tau)\ , \tag{5.37}$$

where we again have made use of the asymptotic formula (5.32) and the approximation $2\sinh\tau\simeq e^\tau$, valid for large τ. Figure 2 shows the logarithm of $\bar{X}=e^{\tau/2}X$ as a function of τ for several values of $\gamma>\gamma_c$. At small τ, where $X\simeq\frac{1}{2}m(m-1)\lambda u^2(\lambda)\simeq$const, $\bar{X}$ increases as $e^{\tau/2}$, while at $\tau\gtrsim\tau_c$, $X\simeq(m-1)/\lambda$ so that $\bar{X}\simeq 2(m-1)e^{-\tau/2}$. The peak in $\bar{X}=\bar{X}(\tau)$ occurs somewhere near the crossover point $\tau=\tau_c$. Using this observation in Eq. (5.37) we find that X_ρ first vanishes at a point ρ_0' roughly determined by the condition $\rho_0'\tau_c\simeq\pi$. Comparison with Eq. (5.35) then shows that $\rho_0=\rho_0'$, i.e., $\rho_0<\rho_1$, and it is the zero ρ_0 (rather than ρ_1) that is the correct counterpart to ρ_0'. From the matching condition $\rho_0\equiv\rho_0'$ now follows the important conclusion that τ_c behaves as $\tau_c\simeq\pi/\rho_0\simeq c(\gamma-\gamma_c)^{-1/2}$ near the critical point. This confirms an assumption basic to the present analysis, namely that the crossover point $\Lambda=\cosh\tau_c$ moves toward infinity as $\gamma\to\gamma_c$.

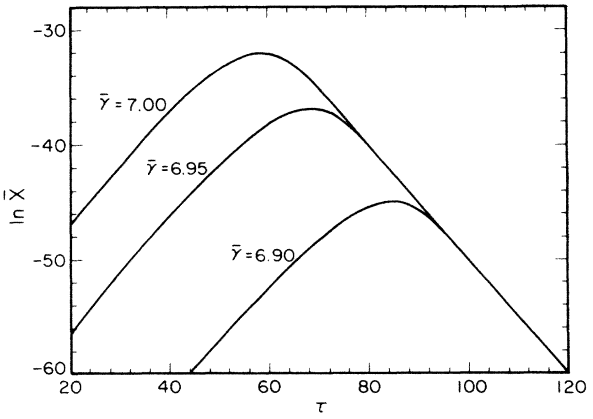

FIG. 2. The logarithm of the quantity $\bar{X}=e^{\tau/2}X$ as a function of the variable $\tau=\mathrm{arccosh}\lambda_1$ for $m=2$.

Efetov has claimed[10] that the appearance of long-ranged solutions u is incompatible with the asymptotic properties of the integral equation (5.19). Roughly speaking, he argues that, after a suitable rescaling, the lhs would go to zero while the rhs would remain finite for $\rho=0$. His argument is false, however, as I now proceed to demonstrate.

For large τ, the kernel K in Eq. (5.25b) integrated over ϕ' depends only on the *difference* $\tau-\tau'$ (or, equivalently, on the ratio λ/λ'), a fact that was already used in the analysis of the localized states in Sec. V A. Efetov argues that due to this asymptotic "translational" invariance of K, the actual solution u to Eq. (5.25a) can be obtained by translation of a special function u_S that solves this equation in the asymptotic domain $\tau\gg1$:

$$u(\cosh\tau)=e^{-\tau_c}u_S(\tau-\tau_c)\ . \tag{5.38}$$

[The prefactor $e^{-\tau_c}$ is required by the conditions $u(\lambda\to\infty)\simeq\lambda^{-1}\simeq 2e^{-\tau}$ and $u_S(\tau\to\infty)\simeq 2e^{-\tau}$.] This is not at all an absurd idea as is shown by a quick glance at Fig. 1. However, although Efetov states that Eq. (5.38) is valid only at large τ, he fails to observe that *not all τ integrals (or λ integrals) in Eqs. (5.33) and (5.34) are dominated by this asymptotic domain.*

Efetov's argument applies without modification to the rhs of Eq. (5.33) because in this case the dominant contributions to the τ integral do come from the region around $\tau=\tau_c\gg1$, see Fig. 2. Using Eq. (5.21) and the representation (5.38), we see that $X=e^{-\tau_c}X_S(\tau-\tau_c)$, and thus

$$X_{\rho=0}\equiv X_0\simeq\tau_c e^{-\tau/2}c_X\ , \tag{5.39}$$

where c_X is a constant independent of τ_c. The difficult part lies in correctly evaluating the integral

$$u_0=\int_0^\infty d\tau\sinh\tau\mathscr{P}_{-1/2}(\cosh\tau)u(\cosh\tau)\ .$$

[Important contributions to this integral arise from *all* parts of the interval $0\le\tau\lesssim\tau_c$. To see this, recall that the solution well to the left of the crossover point is essentially given by $u\propto\mathscr{P}_{-1/2}(\cosh\tau)\simeq\frac{2}{\pi}\tau e^{-\tau/2}$. The exponential factors from $\sinh\tau$, $\mathscr{P}_{-1/2}(\cosh\tau)$ and $u(\cosh\tau)$ thus cancel to produce an integral of the type $\int\tau^2d\tau$.] Equation (5.35) is accurate enough for the present purpose but we still need to determine the constant $u(1)$. We do this

by connecting the solution in the region $\tau \lesssim \tau_c$ with the asymptotic region $\tau \gtrsim \tau_c$. The dominant variation of u in the range $1 \ll \tau \lesssim \tau_c$ comes from the exponential factor $e^{-\tau/2}$ in the asymptotic expansion (5.32). Hence we infer from Eq. (5.38) that

$$u(1) \simeq \text{const} \times e^{-\tau_c/2} , \tag{5.40}$$

which yields

$$u_0 \simeq c_u e^{-\tau_c/2} \int_0^{\tau_c} d\tau \tau^2 = \tau_c^3 e^{-\tau_c/2} c_u/3 . \tag{5.41}$$

Multiplying u_0 with $m - \kappa_0^{-1} \simeq \bar{b}_0(\gamma - \gamma_c)$ and using that τ_c behaves as $\tau_c \propto (\gamma - \gamma_c)^{-1/2}$ we find that the lhs of Eq. (5.33) is given for $\rho = 0$ by $\tau_c e^{-\tau_c/2}$ times a constant.

We thus see that there exists no conflict between the lhs and the rhs of Eq. (5.33) and, in fact, both sides behave in the same way as $\gamma \to \gamma_c$ and $\tau_c \to \infty$. This shows the present analysis to be consistent and justifies *a posteriori* the various assumptions made earlier.

It is now straightforward to derive the critical behavior of the diagonal part of the density-density correlation function, $\Gamma^{\text{ext}}(0;\gamma)$. We recall the relation $Y = 1 - u\lambda$ and use that

$$\begin{aligned} \Gamma^{\text{ext}}(0;\gamma) &= \int_1^{\infty} d\lambda_1 \int_{-1}^{+1} d\lambda_2 Y^{m+1}(\lambda_1,\lambda_2) \\ &\simeq \int_0^{\infty} d\tau e^{\tau}(1 - ue^{\tau}/2)^{m+1} . \end{aligned} \tag{5.42}$$

The second integrand in Eq. (5.42) peaks at $\tau \simeq \tau_c$. Moreover, since $Y = 1 - u\lambda \simeq 1 - e^{\tau - \tau_c} u_S(\tau - \tau_c)/2$, Y assumes a definite shape in the region around $\tau = \tau_c$ which moves toward infinity as $\gamma \to \gamma_c$. We therefore put $Y = Y_S(\tau - \tau_c)$ and write

$$\Gamma^{\text{ext}}(0;\gamma) = e^{\tau_c} \int_0^{\infty} d\tau e^{\tau - \tau_c} Y_S^{m+1}(\tau - \tau_c) . \tag{5.43}$$

The term τ_c appearing in the integrand can be absorbed by shifting the integration variable. Hence, $\Gamma^{\text{ext}}(0;\gamma) \simeq \text{const} \times e^{\tau_c}$, and together with $\tau_c \simeq \text{const} \times (\gamma - \gamma_c)^{-1/2}$ we obtain the following result:

$$\Gamma^{\text{ext}}(0;\gamma) \simeq A \exp\{+[B(\gamma - \gamma_c)]^{-1/2}\} , \tag{5.44}$$

where A and B are nonuniversal constants that depend on the value of m. The numerical results for $\Gamma^{\text{ext}}(0;\gamma)$ quoted at the end of Sec. V B 1 are fit well by Eq. (5.44).

3. *Correlation functions*

This subsection has to begin with another apology. Equation (6.23) of Ref. 15 is not quite correct, and the correct expressions are given by

$$\text{trg} QQ' = \text{trg} Q_{11} Q'_{11} + \text{trg} Q_{12} Q'_{21} + \text{trg} Q_{21} Q'_{12} + \text{trg} Q_{22} Q'_{22} \tag{5.45}$$

and

$$\text{trg} Q_{11} Q'_{11} = -(\lambda_1 \lambda'_1 - \lambda_2 \lambda'_2) - \Delta\alpha^* \Delta\alpha (\lambda_1 - \lambda_2)(\lambda'_1 - \lambda'_2) , \tag{5.46a}$$

$$\text{trg} Q_{22} Q'_{22} = -(\lambda_1 \lambda'_1 - \lambda_2 \lambda'_2) + \Delta\beta^* \Delta\beta (\lambda_1 - \lambda_2)(\lambda'_1 - \lambda'_2) , \tag{5.46b}$$

$$\begin{aligned} \text{trg} Q_{12} Q'_{21} = e^{-i(\gamma_1 + \gamma_2)} \{ & [(\mu_1^*)' \mu_1 + \mu_2^* \mu'_2] + \tfrac{1}{2}(\Delta\alpha^* \Delta\alpha - \Delta\beta^* \Delta\beta)[(\mu_1^*)' \mu_1 - \mu_2^* \mu'_2] \\ & - \tfrac{1}{4} \Delta\alpha^* \Delta\alpha \Delta\beta^* \Delta\beta [(\mu_1^*)' \mu_1 + \mu_2^* \mu'_2] + \Delta\alpha^* \Delta\beta (\mu_1^*)' \mu_2^* + \Delta\beta^* \Delta\alpha \mu_1 \mu'_2 \} , \end{aligned} \tag{5.46c}$$

$$\text{trg} Q_{21} Q'_{12} = (\text{trg} Q_{12} Q'_{21})^* , \tag{5.46d}$$

where

$$i\gamma_1 = \tfrac{1}{2}[(\alpha^*)'\alpha - \alpha^* \alpha'], \quad i\gamma_2 = \tfrac{1}{2}[(\beta^*)'\beta - \beta^* \beta'] . \tag{5.47}$$

Equation (6.23) of Ref. 15 differs from Eqs. (5.45), (5.46) by the phase factor $e^{i(\gamma_1 + \gamma_2)}$. The omission of this factor went undetected because it has no effect for localized states, which was all that was considered in this reference.

Using Eqs. (5.45) and (5.46) and the integral theorem (4.3), one shows that the density-density correlation function $\Gamma^{\text{ext}}(x - y;\gamma)$ can be calculated by the following steps:

$$\begin{aligned} \delta y^{(0)}(\lambda_1,\lambda_2) &= Y(\lambda_1,\lambda_2) , \\ \delta y^{(1)}(\lambda_1,\lambda_2) &= (m+1) \int d\lambda'_1 d\lambda'_2 K(\lambda_1,\lambda_2;\lambda'_1,\lambda'_2) \\ &\quad \times Y^{m-1}(\lambda'_1,\lambda'_2) \delta y^{(0)}(\lambda'_1,\lambda'_2) , \\ &\vdots \\ \delta y^{(p)}(\lambda_1,\lambda_2) &= m \int d\lambda'_1 d\lambda'_2 K(\lambda_1,\lambda_2;\lambda'_1,\lambda'_2) \\ &\quad \times Y^{m-1}(\lambda'_1,\lambda'_2) \delta y^{(p-1)}(\lambda'_1,\lambda'_2) , \\ N(p) \Gamma^{\text{ext}}(p;\gamma) &= \int_1^{\infty} d\lambda_1 \int_{-1}^{+1} d\lambda_2 Y^m(\lambda_1,\lambda_2) \delta y^{(p)}(\lambda_1,\lambda_2) . \end{aligned} \tag{5.48}$$

Note that this sequence of integrations is very similar (not surprisingly so) to the one in Eqs. (5.5). As before, the relationship between $\delta y^{(k)}$ and $\delta y^{(k-1)}$ is nothing but the linearized version of the integral equation (5.19). To see this, we write Eq. (5.19) in the form[10]

$$u=\frac{1-Y}{\lambda_1-\lambda_2}$$
$$=\int_1^\infty d\lambda_1' \int_{-1}^{+1} d\lambda_2' K(\lambda_1,\lambda_2;\lambda_1',\lambda_2') \left[\frac{1-Y^m}{\lambda_1'-\lambda_2'}\right], \quad (5.49)$$

put $Y\to Y+\delta Y$, and define $\delta y=\delta Y/(\lambda_1-\lambda_2)$.

The analysis of Eqs. (5.48) parallels the analysis of Eqs. (5.5) in Sec. V A, and I will simply point out the important differences. We are seeking the eigenvalues and eigenfunctions of the linear operator $M(\lambda;\lambda')=mK(\lambda;\lambda')Y^{m-1}(\lambda')$ with $K(\lambda;\lambda')$ obtained by integrating over ϕ' in Eq. (5.25b). (As discussed earlier, we may ignore all dependence on the compact variable λ_2, and I put $\lambda_1\to\lambda$.) Because M is effectively nonzero only in a finite interval ($1\lesssim\lambda\lesssim\Lambda$), the relevant part of its spectrum is now *discrete* instead of continuous. For $\lambda\ll\Lambda$, where $Y^{m-1}\simeq 1$, the eigenfunctions of M, denoted by $Y_n(\lambda)$, must be approximately proportional to the Legendre functions $\mathscr{P}_{-1/2+i\rho}(\lambda)$, and we choose the normalization such that

$$Y_n(\lambda)\simeq\mathscr{P}_{-1/2+i\rho_n}(\lambda) \quad (\lambda\ll\Lambda) . \quad (5.50)$$

The rapid vanishing of Y^{m-1} at $\lambda\simeq\Lambda$ selects as eigenfunctions of M those Legendre functions that have a node near this point. Using the asymptotic formula (5.32), we find that the ρ_n are approximately determined by the "quantization rule"

$$0=\sin(\rho_n\tau_c)\Longrightarrow\rho_n=\frac{n\pi}{\tau_c}, \quad n=1,2,\ldots . \quad (5.51)$$

The same steps that were performed in Sec. V A now yield ($|x|\gg 1$):

$$N(|x|)\Gamma^{\rm ext}(x;\gamma)\simeq\text{const}\times e^{\tau_c}\tau_c^{-3}\sum_{n=1}^{\infty} n^2(m\bar\kappa_n)^{|x|} , \quad (5.52)$$

where $m\bar\kappa_n$ is the eigenvalue of M belonging to the eigenfunction Y_n. We will need to know the value of the largest eigenvalue, $m\bar\kappa_1$. From the numerical stability of the integral equation with respect to small perturbations, we infer that $m\bar\kappa_1<1$. A lowest-order analytical estimate based on the matching condition discussed below Eq. (5.37) and the quantization rule (5.51) gives $m\bar\kappa_1=1$, which means that for this quantity one has to work more accurately than was done in these equations. Instead of attempting to control the corrections analytically, I have convinced myself by numerical diagonalization of M that indeed $m\bar\kappa_1<1$ and $\lim_{\gamma\to\gamma_c} m\bar\kappa_1=1$. For $m=2$ and $10^2\times\gamma=7.10$, 7.05, 7.00, 6.95, and 6.90, I obtain $-\ln m\bar\kappa_1=1.1\times10^{-3}$, 8.1×10^{-4}, 5.8×10^{-4}, 3.7×10^{-4}, and 2.0×10^{-4}, with an uncertainty of about 1.0×10^{-5} in all cases. From this and $\gamma_c=6.803\times10^{-2}$, I deduce that

$$\xi\simeq\text{const}\times(\gamma-\gamma_c)^{-\nu}, \quad \nu=1.50\pm0.05 , \quad (5.53)$$

where $\xi^{-1}=-\ln m\bar\kappa_1$. If the exponent ν is really equal to $\frac{3}{2}$ as is suggested by Eq. (5.53), then it should be possible to derive this result by a refinement of the present analysis.

Close to the critical point and for $|x|\ll\xi$, we can make a continuum approximation to the sum in Eq. (5.52) and, consequently,

$$N(|x|)\Gamma^{\rm ext}(x;\gamma)\simeq\text{const}\times e^{\tau_c}|x|^{-3/2} . \quad (5.54)$$

In the opposite limit, $|x|\gg\xi$, only the maximum eigenvalue $m\bar\kappa_1$ contributes to the sum in (5.52), resulting in a purely exponential decay,

$$N(|x|)\Gamma^{\rm ext}(x;\gamma)\simeq\text{const}\times e^{\tau_c}\tau_c^{-3}\exp(-|x|/\xi) . \quad (5.55)$$

This concludes the discussion of the critical behavior of the density-density correlation function.

There is one remaining independent two-point Green's function to evaluate, namely $K^{(2)}(x,y;\gamma,\epsilon=0)$. Unfortunately, for extended states this quantity is much more difficult to deal with than $K^{(1)}$. The reason is that the matrix elements Q_{12}^{lk} and Q_{21}^{kl} appearing in the definition of $K^{(2)}$ have a more complicated dependence on the Grassmann variables parametrizing the matrices u and v and, therefore, execution of the Grassmann integrations becomes more difficult (not impossible, just messy). However, if the system is close to $\gamma=\gamma_c$, then essentially the same reasoning that led to Eq. (4.9) for localized states leads to simplifications here; i.e., it can be shown that $K^{(2)}(x,y;\gamma,\epsilon=0)$ behaves like $\Gamma^{\rm ext}(x-y;\gamma)$ as long as $|x-y|\ll\xi$. For $|x-y|\gg\xi$, $K^{(2)}$ is expected to be qualitatively more long ranged. Indeed, according to Eqs. (3.1b), (3.3), and (4.11), $\sum_y K^{(2)}(x,y;\gamma,\epsilon=0)$ must diverge, whereas $\sum_y\Gamma^{\rm ext}(x-y;\gamma)$ should be finite on physical grounds.

VI. DISCUSSION

In this paper, I have performed an extensive numerical and analytical study of the two-point Green's functions for the graded pseudounitary nonlinear σ model on the Bethe lattice. It was shown that at the critical point $\gamma=\gamma_c$ these functions assume a scaling behavior which is of a form standard in critical phenomena. It came as a surprise at first that $\Gamma^{\rm loc}(0;\gamma)$ vanishes discontinuously at $\gamma=\gamma_c$, but it is clear now that this is consistent with the power-law exponent of $\frac{3}{2}$ for the corresponding correlation function. What may seem more surprising is that $\Gamma^{\rm ext}$ has an exponential singularity. This finding is at odds with results presented by Efetov, who has suggested[10] that all thermodynamic observables should be finite at $\gamma=\gamma_c$. According to the present analysis, there definitely is no such thing as a minimum metallic conductivity on the Bethe lattice, although, in Efetov's terminology, there does exist a "maximum dielectric constant."

It has been asserted in Sec. V that the noncompact sector $\mathrm{SU}(1,1)/\mathrm{U}(1)\subset\mathrm{U}(1,1/2)/\mathrm{U}(1/1)\otimes\mathrm{U}(1/1)$ is of overriding importance in determining the nature of the phase transition on the Bethe lattice. Superficially seen, the function of the compact sector and the Grassmann variables is simply to ensure the normalization of certain in-

tegrals. Is that their only function, or do they actually play a more subtle role? The following paragraphs provide a partial answer to this question.

Let us first consider the model which is obtained by restricting the definitions given in Sec. II to the noncompact sector, SU(1,1)/U(1). On the Bethe lattice, the solution of such a model again reduces to the solution of an integral equation of the form (5.1) but now with $Q \in \mathrm{SU}(1,1)/\mathrm{U}(1)$. [Care must be taken to insert a normalization factor Z^{-1} which ensures that the solution Y obeys $\int d\mu(Q) Y(Q) = 1$.] A little thought shows that we are running into a difficulty: the model cannot support localized states. To understand this, note that the kernel of the integral equation now has the asymptotic form

$$\int_0^{2\pi} \frac{d\phi'}{2\pi} L(Q;Q') \simeq l(\lambda/\lambda')/\sqrt{\lambda\lambda'} \quad (\lambda \gg 1) , \tag{6.1}$$

where l has the property $l(z) = l(z^{-1})$. The eigenfunctions of L are the Legendre functions $\mathscr{P}_\nu$, which behave asymptotically as $\sim \lambda^\nu$ ($\mathrm{Re}\nu > -\frac{1}{2}$). For real ν, the minimum eigenvalue of L occurs at $\nu = -\frac{1}{2}$ due to the inverse square root in Eq. (6.1), and the eigenvalues *increase monotonically* with $|\nu + \frac{1}{2}|$. This has the disastrous consequence that the constant (symmetry-unbroken) solution can *never* be stable against the perturbation induced by the symmetry-breaking term $D(Q) = \exp(-2\epsilon\lambda)$. [The culprit is the missing factor $(\lambda_1 - \lambda_2)(\lambda_1' - \lambda_2') \simeq \lambda\lambda'$ which was generated in Sec. V by integration over the Grassmann variables; see Eq. (5.22).] Hence, we are forced to conclude that the system is in the symmetry-broken phase for *all* values of the coupling constant.

Since the noncompact space SU(1,1)/U(1) by itself cannot support a symmetry-unbroken phase, which important feature is it that has been left out? Is it the graded symmetry? To obtain further insight, I have attempted to *grade* the coset space SU(1,1)/U(1). This can be done by taking a complex bosonic variable a ($-\infty < \mathrm{Re}a$, $\mathrm{Im}a < +\infty$) and supplementing it with a complex fermionic variable η to form the graded vector

$$v = \begin{bmatrix} a \\ \eta \end{bmatrix}, \quad v^\dagger = (a^* \ \ \bar{\eta}) . \tag{6.2}$$

From this vector, one then constructs the graded 3×3 matrix

$$Q = \begin{bmatrix} -i(1 + vv^\dagger)^{1/2} & v \\ v^\dagger & +i(1 + v^\dagger v)^{1/2} \end{bmatrix} ,$$

i.e., (6.3)

$$Q \in \mathrm{U}(1,1/1)/\mathrm{U}(1/1) \otimes \mathrm{U}(1) .$$

This model is still much easier to deal with than the original model defined in Sec. II because of the reduced matrix dimension and because it contains only a *single* pair of fermions. Unfortunately, the model (6.3) is not satisfactory either. The minimum eigenvalue of the corresponding kernel $L(Q;Q')$ now lies at $\nu = 0$. While this suffices to stabilize the constant solution for $m = 1$ (one dimension), it does not for a general value of m. In order to move the minimum eigenvalue to positive values of ν and thereby achieve stability, more degrees of freedom must be added. If we do this in the form of a second graded vector and take its bosonic component to the compact (which seems to be forced by convergence and symmetry arguments), then we are led back to the graded pseudounitary nonlinear σ model studied in the present paper.

ACKNOWLEDGMENTS

I thank I. Affleck, H. Nishioka, D. J. Thouless, J. J. M. Verbaarschot, and H. A. Weidenmüller for useful discussions. I would like to thank David Thouless in particular for a critical remark which stimulated this further investigation of the Bethe lattice problem. This work was supported in part by the National Science Foundation, Grants Nos. PHY82-07332 and PHY85-05682.

APPENDIX: ASYMPTOTIC BEHAVIOR OF CONE FUNCTIONS

This appendix contains a derivation of the asymptotic form (5.32) of the functions $\mathscr{P}_{-1/2+i\rho}(\cosh\tau)$ in the limit $\tau \gg 1$, $\rho \ll 1$. Heavy use of this asymptotic form was made in the analysis of Sec. V.

By using the identities

$$a^{-1/2+i\rho} = \frac{1}{\Gamma(\frac{1}{2} - i\rho)} \int_0^\infty \frac{dz}{z} z^{1/2-i\rho} e^{-az} , \tag{A1}$$

$$\int_0^{2\pi} \frac{d\theta}{2\pi} e^{a \sinh\tau \cos\theta} = I_0(a \sinh\tau) , \tag{A2}$$

(I_0 is the modified Bessel function of zeroth order), the integral representation (5.31) for $\mathscr{P}_{-1/2+i\rho}(\cosh\tau)$ can be cast in the form

$$\mathscr{P}_{-1/2+i\rho}(\cosh\tau) = \frac{1}{\Gamma(\frac{1}{2} - i\rho)} \int_0^\infty \frac{dz}{z} z^{1/2-i\rho} e^{-z\cosh\tau} \times I_0(z \sinh\tau) . \tag{A3}$$

As we are interested in the behavior of $\mathscr{P}_{-1/2+i\rho}(\cosh\tau)$ for large τ, it is natural to seek to replace $I_0(\beta)$ by its asymptotic form for large values of the argument,

$$I_0(\beta) \simeq \frac{e^\beta}{\sqrt{2\pi\beta}} \quad (\beta \gg 1) . \tag{A4}$$

But before we can do so, it is necessary to regularize the z integral the origin; otherwise the term $1/\sqrt{\beta} \propto 1/\sqrt{z}$ introduces a spurious singularity at precisely this point. Anticipating the final result (5.32), we place a cutoff on the z integration at $z = e^{-\tau}$. The error caused by this modification of the integral is estimated as

$$\sim \frac{1}{\Gamma(\frac{1}{2} - i\rho)} \int_0^{e^{-\tau}} \frac{dz}{z} z^{1/2-i\rho} = \frac{1}{\Gamma(\frac{1}{2} - i\rho)} \frac{e^{-(\frac{1}{2} - i\rho)\tau}}{(\frac{1}{2} - i\rho)} , \tag{A5}$$

which for $\rho \ll 1$ and most values for τ is much smaller than the expression (5.32).

We now use the asymptotic formula (A4) and the ap-

proximation $\Gamma(\frac{1}{2}-i\rho)\simeq\sqrt{\pi}$, valid for $\rho \ll 1$, to obtain

$$\mathscr{P}_{-1/2+i\rho}(\cosh\tau)\simeq\frac{1}{\sqrt{\pi}}\int_{e^{-\tau}}^{\infty}\frac{dz}{z}z^{1/2-i\rho}\frac{e^{-ze^{-\tau}}}{\sqrt{2\pi z\sinh\tau}}$$
$$\simeq\frac{1}{\pi}e^{-(1/2+i\rho)\tau}\int_{e^{-2\tau}}^{\infty}\frac{dz}{z}z^{-i\rho}e^{-z}. \tag{A6}$$

The final step is to make the substitution $z=e^{w}$, to restrict the w integration to the range $-2\tau\le w\le 0$, and to omit the factor e^{-z} from the integrand. The last two manipulations are justified because, by a similar reasoning as was used above, contributions to the z integral from $z\gtrsim 1$ are small compared to the result (5.32). This yields

$$\mathscr{P}_{-1/2+i\rho}(\cosh\tau)\simeq\frac{1}{\pi}e^{-\tau/2}\int_{-2\tau}^{0}dw\,e^{-i\rho(w+\tau)}, \tag{A7}$$

which equals the expression given in the text.

[1]F. Wegner, Z. Phys. B **35**, 207 (1979).

[2]E. Brezin, S. Hikami, and J. Zinn-Justin, Nucl. Phys. **B165**, 528 (1980).

[3]K. B. Efetov, A. I. Larkin and D. E. Khmel'nitskiĭ, Zh. Eksp. Teor. Fiz. **79**, 1120 (1980) [Sov. Phys.—JETP **52**, 568 (1980).

[4]A. A. M. Pruisken, Nucl. Phys. **B235**, 277 (1984).

[5]H. Levine, S. Libby, and A. A. M. Pruisken, Nucl. Phys. **B240**, 30 (1985); **B240**, 49 (1985); **B240**, 71 (1985).

[6]I. Affleck, Nucl. Phys. **B265**, 409 (1986).

[7]K. B. Efetov, Adv. Phys. **32**, 53 (1983).

[8]J. J. M. Verbaarschot and M. R. Zirnbauer, J. Phys. A **17**, 1093 (1985).

[9]K. B. Efetov, Pis'ma Zh. Eksp. Teor. Fiz. **40**, 17 (1984) [JETP Lett. **40**, 738 (1984)].

[10]K. B. Efetov, Zh. Eksp. Teor. Fiz. **88**, 1032 (1985) [Sov. Phys.—JETP **61**, 606 (1985)].

[11]E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979).

[12]K. B. Efetov, Pis'ma Zh. Eksp. Teor. Fiz. **41**, 508 (1985) [JETP Lett. **41**, 619 (1985)].

[13]R. Abou-Chacra, P. W. Anderson, and D. J. Thouless, J. Phys. C **6**, 1734 (1973).

[14]H. Kunz and B. Souillard, J. Phys. (Paris) Lett. **44**, 411 (1983).

[15]M. R. Zirnbauer, Nucl. Phys. **B265**, [FS15] 375 (1986).

[16]F. J. Wegner, Phys. Rev. B **19**, 783 (1979).

[17]J. J. M. Verbaarschot, H. A. Weidenüller, and M. R. Zirnbauer, Phys. Rep. **129**, 367 (1985).

[18]F. Wegner, Z. Phys. B **36**, 209 (1980).

[19]In the *Note Added in Proof* of Ref. 15 it was falsely claimed that the power-law exponent is given by -2.

[20]N. J. Vilenkin, *Special Functions and the Theory of Group Representations* (AMS Translation, Providence, 1968).

[21]W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Springer, New York, 1966).