Nonperturbative Effects of a Topological Theta Term on Principal Chiral Nonlinear Sigma Models in 2+1 Dimensions

Cenke Xu and Andreas W. W. Ludwig

Phys. Rev. Lett. **110**, 200405 — Published 17 May 2013

DOI: 10.1103/PhysRevLett.110.200405

Nonperturbative effects of a Topological Θ-term on Principal Chiral Nonlinear Sigma Models in (2+1) dimensions

Cenke Xu[1] and Andreas W. W. Ludwig[1]

[1]*Department of Physics, University of California, Santa Barbara, CA 93106*

(Dated: April 24, 2013)

We study the effects of a topological Θ-term on 2+1 dimensional principal chiral models (PCM), which are nonlinear sigma models defined on Lie group manifolds. We find that when $\Theta = \pi$, the nature of the disordered phase of the principal chiral model is strongly affected by the topological term: it is either a gapless conformal field theory, or it is gapped and two-fold degenerate. The result of our paper can be used to analyze the boundary states of three dimensional symmetry protected topological phases.

PACS numbers:

Introduction: It is well known that topological terms in field theories are responsible for many profound phenomena in condensed matter physics. For instance, the one-dimensional SU(2) spin-1/2 Heisenberg quantum spin chain is known to be described by the (1+1)-dimensional O(3) Nonlinear Sigma Model [(1+1)d NLSM] with a Θ-term, for a (real) three-component unit vector $\vec{n}$ on the two-dimensional sphere S^2 [1–3] with action

$$S = \int d^2x \; \frac{1}{g}(\partial_\mu \vec{n})^2 + \frac{i\Theta}{8\pi}\epsilon_{\mu\nu}\epsilon_{abc}\; n^a \partial_\mu n^b \partial_\nu n^c. \quad (1)$$

Throughout this article we will always work in *imaginary time*. The Θ-term contributes a factor $\exp(i\Theta q)$ to the partition function of the NLSM for every field configuration $\vec{n}(x)$ in (1+1)d space-time which has instantons of 'topological charge' q. For a spin-s chain $\Theta = 2\pi s$, which leads to the qualitative difference between integer and half-integer spin chains, due to the constructive and destructive interference between even and odd number of instantons for the two different values of Θ. A similar (2+0)-d NLSM with a Θ-term can be used to describe the integer quantum Hall plateau and the transition between such plateaus (see also the Supplementary Material).

In the present article, we will study the (2+1)-d Principal Chiral Nonlinear Sigma Model (PCM) with a theta term, which has the action

$$\begin{aligned} S \; &= \int d\tau d^2x \; \frac{1}{g}\mathrm{tr}[\partial_\mu U^\dagger \partial_\mu U] \\ &+ \; \frac{i\Theta}{24\pi^2}\epsilon_{\mu\nu\rho}\mathrm{tr}[(U^\dagger\partial_\mu U)(U^\dagger\partial_\nu U)(U^\dagger\partial_\rho U)]. \end{aligned} \quad (2)$$

where U is a group element that belongs to a (simple) compact Lie group G, such as SU(N), SO(N), Sp(N). All these groups have nontrivial homotopy group $\pi_3[G] = Z$, which implies that the corresponding PCMs possess instantons in (2+1)-d space-time, and that a Θ-term can be added to the action (as in Eq. 2). For arbitrary values of Θ the PCM in Eq. 2 is invariant under a $G_L \times G_R$ symmetry, denoting left and right multiplication of U by group elements. When $\Theta = \pi k$ with integer k, the system also has other discrete symmetries such as reflection $x \to -x$, or time-reversal that transforms $i \to -i$ (we assume U carries a trivial representation of time-reversal). Thus these discrete symmetries guarantee that when $\Theta = \pi k$, Θ does not flow under the renormalization group (RG), while for any other value Θ is not forbidden to flow.

Very recently the new concept of symmetry protected topological (SPT) phases appeared, which has attracted significant attention [4–13]. The PCM in Eq. 2 has been used as a general formalism to describe SPT phases [4, 6, 9, 11–13]. The quantum spin Hall (QSH) insulator in two-, and the Z_2- topological insulator in three spatial dimensions are both well-known examples of SPT phases. In the supplementary material of this paper we will argue that the PCM in Eq. 2 with SU(2) manifold and $\Theta = 2\pi$ is the effective field theory for the (2+1)d QSH insulator with interaction.

Below, we will present our results first for the special case of $G = \mathrm{SU}(2)$. Subsequently, we will explain that our arguments and conclusions are in fact generally applicable to any (simple) compact Lie group G. In the case of $G = SU(2)$ one can parametrize any group element in terms of a four-dimensional (real) unit vector $\vec{\phi}^t = (\phi^0, \phi^1, \phi^2, \phi^3)$ on the 3-dimensional sphere S^3 as $U = \phi^0 + i\phi^1\sigma^x + i\phi^2\sigma^y + i\phi^3\sigma^z$. Now, the $G_L \times G_R = SU(2)_L \times SU(2)_R$ symmetry group is isomorphic to $SO(4)$. This implies that for $G = SU(2)$, the PCM in Eq. 2 is equivalent to the (2+1)-d O(4) NLSM with action

$$S = \int d^3x \; \frac{1}{g}(\partial_\mu\vec{\phi})^2 + \frac{i\Theta}{12\pi^2}\epsilon_{\mu\nu\rho}\epsilon_{abcd}\; \phi^a\partial_\mu\phi^b\partial_\nu\phi^c\partial_\rho\phi^d. \quad (3)$$

The goal of this paper is to provide a non-perturbative argument for the phase diagram of the (2+1)-d PCM with Θ−term, Eq. 2, in terms of the two coupling constants g and Θ. Our result for this phase diagram is depicted in Fig. 2 below. Note that in the absence of space-time boundaries we are allowed to compactify (2+1)-d space-time into a three dimensional sphere S^3. First consider the case where Θ is an integer multiple of 2π, $\Theta = 2\pi k$, so that the Θ-term contributes a factor of unity to the partition function for any nontrivial instanton configuration in the space-time. Thus, in the absence of boundaries in space-time the phase diagram of PCM in Eq. 3 at $\Theta = 2\pi k$ is identical to the model at $\Theta = 0$. For small values of the coupling g, the system is in an O(4) ordered phase with nonzero order parameter $\langle \vec{\phi} \rangle \neq 0$ and three gapless Goldstone modes. For large g, on the other hand, the system is in a quantum disordered phase with a non-degenerate ground state and a fully gapped spectrum. The quantum phase transition between the O(4) ordered and disordered phases is an ordinary 2nd order transition in the 3D O(4) Wilson-Fisher (WF) universality class.

The presence of a Θ-term is not expected to affect the ordered phase, because in it instantons are suppressed. Therefore, it can only play a role for the transition into the disordered phase and in the disordered phase itself. In order to understand the disordered phase and the phase transition in the presence of a $\Theta-$term, standard perturbative methods fail. Thus in order to understand the disordered phase of the PCM in Eq. 3, a non-perturbative argument must be developed, which is what we do in this article. Our conclusion is that there are two possibilities for the disordered phase of the PCM with a $\Theta-$term when $\Theta = \pi$ (Eq. 3 and Eq. 2): It is either a gapless phase with power-law correlations for the fields U (or $\vec{\phi}$), or it is a gapped phase but with a two-fold ground state degeneracy.

O(3) NLSM at $\Theta = \pi$ in (1+1) dimensions: Before we start our argument in (2+1)-d, let us first consider the (1+1)-d O(3) NLSM with the Θ-term in Eq. 1, and focus on $\Theta = \pi$. Since this model describes the SU(2) spin-1/2 chain, we know from the Lieb-Schultz-Mattis (LSM) theorem [15] that this NLSM is either a gapless conformal field theory (CFT), or gapped but two-fold degenerate. Since the goal of this paper is to understand the (2+1)-d PCM field theories of Eq. 2 and Eq. 3 without recourse to any lattice spin model representations, we will first use a new argument to understand the behavior of the (1+1)-d O(3) NLSM without using any lattice representations such as spin chains. Subsequently, we will generalize this argument to the (2+1)-d models.

Our argument proceeds in four steps:

Step (1). In order to understand the O(3) NLSM Eq. 1 at $\Theta = \pi$, let us first look at $\Theta = 0$ and $\Theta = 2\pi$. At these values of Θ, the system also has the discrete symmetry $\vec{n} \to -\vec{n}$, in addition to the SO(3) rotation symmetry. The bulk spectra for $\Theta = 0$ and $\Theta = 2\pi$ are identical, possessing a non-degenerate ground state and a gap to all excitations.

Step (2). Now let us consider the system on a spatial interval with open boundaries at $x = 0$ and $x = L$. Although the models with $\Theta = 0$ and $\Theta = 2\pi$ have identical bulk spectra, they behave very differently at the boundaries. Since the bulk is gapped, we can safely ignore the bulk, and focus on the boundary because the gap in the bulk will protect the effective boundary theory from any singular contributions. Since the boundary is a point in space, it is effectively described by a (0+1)-d O(3) NLSM model. When $\Theta = 0$ this (0+1)-d NLSM model is completely trivial. However, when $\Theta = 2\pi$, the Θ-term in Eq. 1 can be viewed as the O(3) Wess-Zumino-Witten (WZW) term for the (0+1)-d O(3) NLSM model at each of the two boundaries, at $x = 0$ and at L:

$$\int_0^L dx \int d\tau \frac{i2\pi}{8\pi} \epsilon_{\mu\nu}\epsilon_{abc}\ n^a \partial_\mu n^b \partial_\nu n^c = \mathrm{WZW}_0 - \mathrm{WZW}_L. \tag{4}$$

The WZW term for a 0+1 dimensional O(3) NLSM, appearing on the right hand side, is defined as follows: In the (0+1)-d NLSM the O(3) vector $\vec{n}$ is a function only of imaginary time τ. Consider a periodic evolution of $\vec{n}(\tau)$, namely $\vec{n}_{\tau=0} = \vec{n}_{\tau=\beta}$. Then $\vec{n}$ is a mapping from a closed loop S^1 parametrized by $\tau \in [0, \beta]$ to the target space S^2. The WZW term is defined as the solid angle on the target space S^2 enclosed by the closed loop $\tau \in [0, \beta]$. The WZW term at level k can be explicitly written as

$$\mathrm{WZW} \ = \ 2\pi \int d\tau \int_0^1 du\ \frac{ik}{8\pi} \epsilon_{abc}\epsilon_{\mu\nu} n^a \partial_\mu n^b \partial_\nu n^c. \tag{5}$$

Here, the function $\vec{n}(\tau)$ has been extended to a mapping $\vec{n}(\tau, u)$ from a disc (τ, u), where $0 \leq u \leq 1$ and $\tau \in S^1$, to the target S^2. This extension has one constraint: $\vec{n}(\tau, 1) = \vec{n}(\tau)$, $\vec{n}(\tau, 0) = \hat{z}$. Unlike the Θ-term, the coefficient k in Eq. 5 has to be an integer[14], regardless of whether any discrete symmetry is present or not. By simply identifying u with x, one arrives at Eq. 4.

It is well known that if a (0+1)-d O(3) NLSM, describing the quantum mechanics of a point particle on a sphere, has a WZW term at level k, the ground state of this quantum mechanics is $(k + 1)$-fold degenerate. In fact, the ground state of the (0+1)-d O(3) NLSM with a WZW term at level k precisely describes a single SU(2) spin with $S = k/2$. In Eq. 4, the WZW term at each boundary is at level $k = 1$. This implies that the model in Eq. 1 has two fold degeneracy at each boundary when $\Theta = 2\pi$. This conclusion again agrees with Haldane's conjecture, which states that the model with $\Theta = 2\pi$ describes the spin-1 chain. Moreover it recovers the well-known fact that the Haldane phase of the spin-1 chain has an unpaired spin-1/2 degree of freedom at each of its boundaries [16–18].

Step (3). Now let us tune Θ in Eq. 1 continuously from 2π to 0. Then the spin-1/2 boundary state has to disappear at a certain value of Θ. When Θ is tuned away from 2π, the discrete symmetry $\vec{n} \to -\vec{n}$ of the system is broken. One important fact is that the spin-1/2 boundary state cannot be destroyed without going through a bulk transition, even when the discrete symmetry is broken. This is because, given a single spin-1/2, as long as

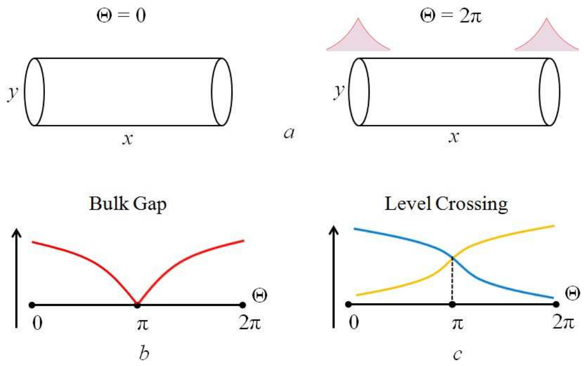

FIG. 1: (a). We compactify the space of model Eq. 2 and Eq. 3 to a two dimensional cylinder. When $\Theta = 2\pi$ there are gapless boundary states localized at the two boundaries. (b). The first possibility when we tune Θ from 2π to 0, the bulk gap closes at $\Theta = \pi$. (c). The second possibility, the two states at $\Theta = 0$ and $\Theta = 2\pi$ have level crossing at $\Theta = \pi$.

the SO(3) symmetry is preserved, the spin-1/2 doublet degeneracy cannot be lifted. This conclusion can also be drawn by noticing that the coefficient of the WZW term has to be quantized, no matter whether the discrete symmetry is broken or not.

$\Theta = k\pi$ with integer k is a fixed point under renormalization group, while with any other value Θ in principle can flow under RG. We should use the fixed point values of Θ to derive the edge WZW model. The paragraph above implies that when $\Theta = 2\pi \pm \epsilon$, the edge state is identical with the edge state with $\Theta = 2\pi$. If one introduces the standard CP^1 representation like in Ref. [18], the $\Theta-$term in Eq. 1 becomes the $\Theta-$term of the U(1) gauge field, and when $\Theta \in (-\pi, 3\pi)$ there is always a gauge charge-1 (spin-1/2) localized at the boundary [22].

Step (4). We have concluded that in order to destroy the boundary spin-1/2 state, a bulk transition has to occur. Here we assume the simplest case, *i.e.* that there is one single transition between $\Theta = 0$ and 2π. Then in this case there are exactly two possibilities for this bulk transition:

4A. The transition is of second order, meaning that the bulk gap closes continuously for some values of Θ between $\Theta = 0$ and 2π. Because the bulk spectrum is identical for Θ and $(2\pi - \Theta)$, this transition has to occur at $\Theta = \pi$ if the bulk gap closes only at *one* value of Θ. This implies that the bulk is gapless when $\Theta = \pi$. When Θ is approaching π from 2π, the boundary spin-1/2 state will become more and more delocalized, and eventually gets absorbed by the gapless bulk states at $\Theta = \pi$.

4B. The transition is of first order, meaning there is always a bulk gap. However, at this first order transition, the two phases with $\Theta = 0$ and $\Theta = 2\pi$ will have a crossing of their ground state energies, and this level crossing also has to occur at $\Theta = \pi$. This implies that the ground state at $\Theta = \pi$ is two-fold degenerate. In this case, when Θ is tuned from 2π to π, the boundary spin-1/2 states will never delocalize, they will simply disappear abruptly at $\Theta = \pi$. An example of this phenomenon is the (1+1)-d CP^N model which is known[21–23] to have a first order transition at $\Theta = \pi$ when $N \geq 3$.

These two possibilities that we have arrived at above are completely consistent with the conclusion drawn from the LSM theorem for the SU(2) spin-1/2 chain.

<u>*O(4) NLSM at $\Theta = \pi$ in (2+1) dimensions:*</u> We will now generalize the arguments given in the previous paragraph to the (2+1)-d O(4) NLSM with a Θ-term, Eq. 3. Since we are only interested in the nature of the disordered phase of Eq. 3, we will consider the case with large values of the coupling g.

Step (1). In order to investigate the disordered phase at $\Theta = \pi$, we first look, as before, at $\Theta = 0$ and 2π. Again, in the bulk these two disordered phases are both gapped with a non-degenerate ground state, while they have different boundary states. In order to look at the boundary states, we let the x direction be a finite interval $0 \leq x \leq L$ while the y direction is periodic, so that the system is defined on a finite 2d cylinder (Fig. 1a).

Step (2). At each boundary located at $x = 0$ and $x = L$ there is a (1+1)-d theory defined on (y, τ)-spacetime. Since the bulk is gapped when $\Theta = 0$ and 2π, the kinetic term of the effective boundary theory is still that of a local O(4) NLSM, Eq. 3, but now in (1+1)-d. When $\Theta = 0$, there is no nontrivial topological term at the boundary. However, when $\Theta = 2\pi$, the Θ-term of the (2+1)-d bulk $O(4)$ NLSM, Eq. 3, can as before be viewed as a WZW term of the (1+1)-d $O(4)$ NLSM appearing on the boundary. Thus, the boundary theory is described by the following (1+1)-d O(4) NLSM with a WZW term:

$$S = \int dy d\tau \ \frac{1}{g}(\partial_\mu \vec{\phi})^2 + \frac{i2\pi k}{12\pi^2} \int dy d\tau \int_0^1 du \epsilon_{\mu\nu\rho}\epsilon_{abcd} \ \phi^a \partial_\mu \phi^b \partial_\nu \phi^c \partial_\rho \phi^d . \quad (6)$$

When $\Theta = 2\pi$, the WZW term in Eq. 6 has level $k = 1$. It is well known[19, 20] that the long-distance behavior of this (1+1)-d O(4) NLSM with level $k = 1$ WZW term is controlled by a stable fixed point at finite g^*, and that this fixed point is precisely the $\mathrm{SU}(2)_1$ CFT which describes the nearest neighbor spin-1/2 Heisenberg chain. When $\Theta = 2\pi k$ and $k =$ integer, the boundary is described at long scales by the $SU(2)_k$ CFT[19, 20]. Thus once again, when Θ is a nonzero integer multiple of 2π the system possesses nontrivial gapless boundary states.

Step (3). The same strategy that we used before can now be applied: When we tune Θ continuously from 2π to 0, then the boundary state has to disappear through a bulk phase transition. This is because SO(4) symmetry of the (1+1)-d theory in Eq. 6 is, as mentioned in the introduction, isomorphic to the $SU(2)_L \times SU(2)_R$ symmetry of the PCM. It is this symmetry that protects the finite-coupling fixed point at $g = g^*$ of the boundary NLSM, Eq.6, from being gapped out. In order to gap out this fixed point CFT, we need to break the $SU(2)_L \times SU(2)_R$ symmetry down to the diagonal SU(2) symmetry, *i.e.* we need to induce a relevant back-scattering between left and right moving boundary modes. However, since our model Eq. 3 has $O(4) \sim SU(2)_L \times SU(2)_R$ symmetry for any value of the coupling g, such backscattering processes are absent. Thus, although tuning Θ away from 2π breaks a discrete symmetry of the system, the boundary CFT cannot be gapped out without going through a bulk transition.

Step (4). Since the boundary states can only be destroyed through a bulk transition, there are the following two possibilities for this transition:

4A. This bulk transition is of second order, and it has to occur at $\Theta = \pi$. This implies that the disordered phase of the PCM in Eq. 3 is gapless at $\Theta = \pi$. Since the transition is of second order, the second derivative $\partial^2 E(\Theta)/\partial\Theta^2$ of the ground state energy $E(\Theta)$ has a singularity at $\Theta = \pi$ (Fig. 1*b*).

4B. This bulk transition is first order and occurs at $\Theta = \pi$. At this transition the two gapped phases with $\Theta = 0$ and $\Theta = 2\pi$ will have a crossing of their ground state energies, and this level crossing has to appear at $\Theta = \pi$. This implies, as before, a gapped spectrum and a two-fold degenerate ground state at $\Theta = \pi$. In this case, the ground state energy $E(\Theta)$ has a kink at $\Theta = \pi$, *i.e.* the first order derivative $\partial E(\Theta)/\partial\Theta$ is discontinuous at $\Theta = \pi$ (Fig. 1*c*).

PCM on group G at $\Theta = \pi$ in (2+1) dimensions: It is straightforward to generalize these arguments to other (2+1)-d PCMs with a Theta term, as in Eq. 2, defined on more general compact Lie group manifolds such as e.g. $G = SU(N), SO(N)$ and $Sp(N)$. The key argument rests on the gaplessness of the CFT that describes the long-distance behavior of the (1+1)-d PCM with WZW term at level k, which appears at the boundary of the (2+1)-d bulk PCM at $\Theta = 2\pi k$. The gaplessness of this CFT is protected by the $G_L \times G_R$ symmetry of the PCM with WZW term, which forbids[20] all operators that are relevant in the RG sense from appearing in the action.

Based on these arguments we obtain the two possibilities for the RG flow diagram of the two coupling constants g and Θ of model Eq. 2 sketched in Fig. 2.

In summary, using a nonperturbative argument, we conclude that the topological $\Theta-$term can drastically change the dynamics of the quantum disordered phases of PCMs in (2+1)-d space-time, when $\Theta = \pi$. As we mentioned at the beginning of this paper, the notion of a SPT phase has become an important concept in condensed matter theory. The key of constructing a SPT phase is to prove that its boundaries are either gapless or possess a degeneracy. The boundary states of many (3+1)-d SPT phases can be mapped exactly to a (2+1)-d PCMs with $\Theta = \pi$ [10, 11]. Thus the result in the current paper immediately concludes that the boundaries of these SPT phases are nontrivial.

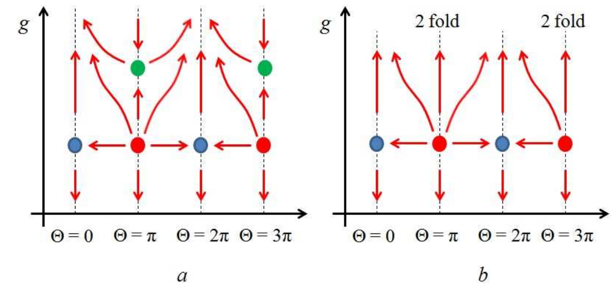

FIG. 2: The two possible RG flows for the coupling constants g and Θ of the (2+1)-d PCM on a compact Lie group G with Theta term, Eq. 2 (and Eq. 3 for $G = SU(2)$).

Acknowledgments.– This work has been supported in part, by NSF DMR-0706140 (A.W.W.L.). C. Xu is supported by the Alfred P. Sloan Foundation, the David and Lucile Packard Foundation, Hellman Family Foundation, and NSF Grant No. DMR-1151208. Both authors acknowledge the hospitality of KITP as organizers of the programs on 'Topological Insulators and Superconductors' (A.W.W.L.) and 'Holographic Dualities (C.X.), and support from the National Science Foundation at under Grant No. NSF PHY05-51164 (KITP).

[1] I. Affleck, Nucl. Phys. B **257**, 397 (1985).
[2] F. D. M. Haldane, Phys. Lett. A **93**, 464 (1983).
[3] F. D. M. Haldane, Phys. Rev. Lett. **50**, 1153 (1983).
[4] Xie Chen, Zheng-Cheng Gu, Zheng-Xin Liu, and Xiao-Gang Wen, arXiv:1106.4772 (2011).
[5] Michael Levin, Ady Stern, Phys. Rev. B 86, 115131 (2012).
[6] Zheng-Xin Liu, Xiao-Gang Wen, arXiv:1205.7024 (2012).
[7] Michael Levin, Zheng-Cheng Gu, Phys. Rev. B 86, 115109 (2012).
[8] Yuan-Ming Lu, Ashvin Vishwanath, Phys. Rev. B 86, 125119 (2012).
[9] T. Senthil, Michael Levin, Phys. Rev. Lett. 110, 046801 (2013).
[10] Ashvin Vishwanath, T. Senthil, arXiv:1209.3058 (2012).
[11] Cenke Xu, arXiv:1209.4399 (2012).
[12] Jeremy Oon, Gil Young Cho, Cenke Xu, arXiv:1212.1726 (2012).
[13] Cenke Xu, T. Senthil arXiv:1301.6172 (2013).
[14] This ensures that two different extensions of $\vec{n}(\tau)$ yield that same contribution to the partition function.
and
[15] E. H. Lieb, T. D. Schultz, and D. C. Mattis, Ann. Phys. **16**, 407 (1961).
[16] M. Hagiwara, K. Katsumata, I. Affleck, B. I. Halperin, and J. P. Renard, Phys. Rev. Lett. **65**, 3181 (1990).
[17] S. H. Glarum, S. Geschwind, K. M. Lee, M. L. Kaplan,

and J. Michel, Phys. Rev. Lett. **67**, 1614 (1991).
[18] T.-K. Ng, Phys. Rev. B **50**, 555 (1994).
[19] E. Witten, Commun. Math. Phys. **92**, 455 (1984).
[20] V.G. Knizhnik and A.B. Zamolodchikov, Nucl. Phys. **B247**, 83 (1984).
[21] E. Witten, Nucl. Phys. B **149**, 285 (1979).
[22] S. Coleman, Ann. of Phys. **101**, 239 (1976).
[23] I. Affleck, Nucl. Phys. B **305**, 582 (1988).