Spectrum of Dirac operator and role of winding number in QCD

H. Leutwyler and A. Smilga*
Institute for Theoretical Physics, University of Bern, Sidlerstrasse 5, CH-3012 Bern, Switzerland
(Received 4 August 1992)

We show that very general considerations based on the properties of the partition function of QCD allow one to extract information about the eigenvalues of the Dirac operator in vacuum gauge fields. In particular, we demonstrate that the familiar suppression of field configurations with a nontrivial topology occurring for small quark masses is a finite size effect which disappears if the four-dimensional volume V is large enough. The formation of a quark condensate is connected with the occurrence of small eigenvalues of order $\lambda_n \propto 1/V$.

PACS number(s): 11.15.Tk, 11.30.Qc, 11.30.Rd, 12.38.Aw

I. INTRODUCTION

The properties of the QCD vacuum are still poorly understood. From numerical computations on a lattice [1] and also from the observed hadron spectrum, analyzed in terms of QCD sum rules [2], we only know its most crude characteristics, such as the gluon condensate $\langle 0|G^a_{\mu\nu}G^a_{\mu\nu}|0\rangle$ and the quark condensate $\langle 0|\bar{q}q|0\rangle$. The latter is of particular interest as it signals spontaneous breaking of the chiral $SU_L(N_f)\times SU_R(N_f)$ symmetry which the QCD Lagrangian enjoys in the massless quark limit.

Possible mechanisms for spontaneous chiral-symmetry breaking and the formation of nonperturbative condensates have repeatedly been discussed in the literature (for a comprehensive review of the subject, see, e.g, [3]), but a full understanding of the problem has not been achieved—a derivation of spontaneous symmetry breakdown from first principles yet remains to be found. Of particular importance is the observation that the fermion condensate is related to the mean density of the fermion eigenvalues. At infinite volume, the relation is the following. Denote by $\rho(\lambda)d\lambda$ the mean number of eigenvalues contained in the interval $d\lambda$, per unit volume. The quark condensate is determined by the level density at $\lambda=0$ [4]:

$$\langle 0|\bar{q}q|0\rangle = -\pi\rho(0) . \tag{1.1}$$

Let us recall how this beautiful relation is derived. Treating the gauge field $G^a_\mu(x)$ as an external field, the fermion Green's function is given by

$$S_G(x,y)=\langle q(x)\bar{q}(y)\rangle_G=\sum_n \frac{u_n(x)u_n^\dagger(y)}{m-i\lambda_n} , \tag{1.2}$$

where $u_n(x)$ and λ_n are eigenfunctions and eigenvalues of the Euclidean Dirac operator:

$$\not{D}u_n(x)=\lambda_n u_n(x) . \tag{1.3}$$

Except for zero modes, the eigenfunctions occur in pairs $u_n, \gamma_5 u_n$ with opposite eigenvalues. Setting $x=y$ and integrating over x, the representation (1.2) therefore implies

$$\frac{1}{V}\int_V dx\,\langle \bar{q}(x)q(x)\rangle_G = -\frac{2m}{V}\sum_{\lambda_n>0}\frac{1}{m^2+\lambda_n^2} , \tag{1.4}$$

where the zero-mode contributions have been dropped (we will see later under what conditions this is justified). The quark condensate is the average of the left-hand side over all gluon-field configurations, and the vacuum expectation value $\langle 0|\bar{q}q|0\rangle$ is the infinite-volume limit thereof. In this limit the level spectrum becomes dense, the mean number of eigenvalues contained in a given interval being proportional to the volume. Averaging the relation (1.4) over all gluon-field configurations and taking the limit $V\to\infty$, we get

$$\langle 0|\bar{q}q|0\rangle = -2m\int_0^\infty d\lambda \frac{\rho(\lambda)}{m^2+\lambda^2} , \tag{1.5}$$

where $\rho(\lambda)$ is the mean level density introduced above.

Actually, the formula (1.5) does not make sense as it stands: Perturbation theory shows that the integral is quadratically divergent at the upper end. The problem is related to the fact that the energy density of the vacuum contains ultraviolet-divergent terms of the form Λ_0^4, $m^2\Lambda_0^2$, and $m^4\ln\Lambda_0$, where Λ_0 is the cutoff. The renormalization of the coupling constant and quark mass does not suffice to ensure that the partition function remains finite if the cutoff is removed. One, in addition, needs to introduce a cosmological constant of the form $c_0+c_1m^2+c_2m^4$ and tune the coefficients c_i. The result for $\langle 0|\bar{q}q|0\rangle$ is unique except for finite contributions proportional to m and to m^3, respectively. In Eq. (1.5) these contributions show up when writing the high-frequency part of the dispersion integral in convergent (twice subtracted) form:

$$\langle 0|\bar{q}q|0\rangle = -2m\int_0^\mu \frac{d\lambda\,\rho(\lambda)}{m^2+\lambda^2} -2m^5\int_\mu^\infty \frac{d\lambda}{\lambda^4}\frac{\rho(\lambda)}{m^2+\lambda^2}+\gamma_1 m+\gamma_2 m^3 . \tag{1.6}$$

*On leave of absence from ITEP, Moscow, Russia.

In the limit $m \to 0$, the entire contribution from the interval $\mu < \lambda < \infty$ disappears and only the infrared part contained in the first term of Eq. (1.6) can generate a nonzero result.

The contribution from small values of λ depends on the behavior of the spectral density there: If $\rho(\lambda) \propto \lambda^\alpha$, the right-hand side (RHS) of Eq. (1.5) is proportional to m^α. Hence a finite result arises if and only if $\rho(\lambda)$ tends to a nonzero limit for $\lambda \to 0$. Performing the integration over λ for this case, we indeed arrive at Eq. (1.1).

In the above consideration, the infinite-volume limit plays an essential role. In fact, at finite volume, the behavior of the quark condensate in the chiral limit for $N_f \geq 2$ is quite different, because symmetries do not break down spontaneously if V is finite. The spectrum is then discrete, and the sum occurring in Eq. (1.4) therefore does not develop an infrared singularity. If the limit $m \to 0$ is taken at fixed volume, the condensate disappears—chiral symmetry is restored. In the above analysis, we first took the limit $V \to \infty$, followed by $m \to 0$, and arrived at a nonzero condensate. Clearly, the two limits are not interchangeable.

Actually, the above analysis does go through also at finite volume, provided only that the quark mass is not too small. To see this we first observe that the mean spacing between the levels is inversely proportional to the volume, $\Delta\lambda \simeq 1/V\rho(\lambda)$. In particular, at the lower end of the spectrum, the relation (1.1) implies $\Delta\lambda \simeq \pi/(V|\langle 0|\bar{q}q|0\rangle|)$. If the quark mass is large compared to the level spacing,

$$Vm|\langle 0|\bar{q}q|0\rangle| \gg 1 , \tag{1.7}$$

the factor $(m^2+\lambda_n^2)^{-1}$ only slowly varies with n, so that it is legitimate to replace the sum in Eq. (1.4) by the corresponding integral. Hence the representation (1.5) also holds at finite volume, provided the mass obeys the condition (1.7). If the volume is large, this condition only excludes academically small quark masses. We conclude that the spontaneous formation of a quark condensate in the chiral limit at infinite volume can also be seen if the volume is finite. It is related to the fact that the Dirac operator admits a discrete spectrum of eigenvalues which are inversely proportional to the volume, $\lambda_n \sim 1/V$. The purpose of the present paper is to study the distribution of these eigenvalues in some detail. In particular, we will investigate the role played by the winding number which characterizes the topology of the gauge field. Nontrivial topologies necessarily give rise to fermionic zero modes which tend to suppress the fermion determinant if the quark masses are small. One of the goals of the present investigation is to determine the relative weight of topologically nontrivial gauge-field configurations, i.e., to calculate the distribution of the winding number.

The paper is organized as follows. In Secs. II and III we discuss the properties of the QCD partition function for finite volume and in the presence of a nonzero vacuum angle θ. We exploit the well-known fact that the full partition function $Z(\theta)$ can be represented as a Fourier series over partition functions Z_ν which correspond to gauge-field configurations of fixed winding number ν. Furthermore, we make use of a general property which holds for any theory with a mass gap in the phyiscal spectrum, such as QCD, with less than two quark flavors: If the size of the box is large compared to the Compton wavelength of the lightest particle, the logarithm of the partition function is an extensive quantity, up to exponentially small finite-size effects.

In Secs. IV–VII we apply these considerations to the partition function of QCD with $N_f=1$. In particular, we show that if the quark mass obeys the condition (1.7), the contributions to the fermion condensate from gauge field configurations with different winding numbers $\nu=0,\pm 1,\pm 2,\ldots,$ are all equally important. Amusing peculiarities occurring if the theory is restricted to topologically trivial configurations are pointed out in Sec. V.

The implications of this analysis for the spectrum of the Dirac operator are discussed in Sec. VI. We show that the occurrence of small eigenvalues is a necessary corollary of the properties established for Z_ν. In Sec. VII we analyze the different contributions to the quark condensate and demonstrate that, for $\theta=0$, the basic relation (1.1) holds if and only if the parameter $Vm|\langle 0|\bar{q}q|0\rangle|$ is large and positive. The case of two or more flavors is discussed in Secs. VIII and IX. The structure of the partition function is more complicated here, because the theory now contains Goldstone bosons. We exploit the fact that if all four sides of the box are large, only Goldstones of small momenta occur. Their properties are fixed by symmetry considerations and their contribution to the partition function can be analyzed by means of an effective Lagrangian. The analysis performed for the case $N_f=1$ then goes through without any essential modifications. In Sec. X we discuss the modifications occurring if the fermions are taken in the adjoint rather than in the fundamental representation of color. In this case gauge-field configurations with fractional winding number play an important role, but otherwise the main results of our analysis remain unaffected. Finally, in Sec. XI we investigate the behavior of the winding-number distribution and eigenvalue spectrum if the number of colors, N_c, is taken as large. Section XII contains a brief summary and some conclusions.

II. PARTITION FUNCTION ON A TORUS

In Euclidean space the Lagrangian is of the form[1]

$$L=\frac{1}{4g^2}G^a_{\mu\nu}G^a_{\mu\nu}-i\theta\omega-i\bar{q}\not{D}q+\bar{q}_R\mathcal{M}q_L+\bar{q}_L\mathcal{M}^\dagger q_R , \tag{2.1}$$

where the Dirac operator $\not{D}=\gamma_\mu(\partial_\mu+i\mathbf{G}_\mu)$ includes the gluon field $\mathbf{G}_\mu(x)$ and $\omega(x)$ is the corresponding winding-number density:

$$\omega=\frac{1}{32\pi^2}G^a_{\mu\nu}\tilde{G}^a_{\mu\nu} . \tag{2.2}$$

With a suitable choice of the quark-field basis, the mass term can be brought to the standard form where the matrix $\mathcal{M}$ is diagonal with real positive entries $m_u, m_d, \ldots$.

[1]We normalize the Euclidean γ matrices by $\{\gamma_\mu,\gamma_\nu\}=-2\delta_{\mu\nu}$ and use a representation where $\not{D}$ is Hermitian.

Consider now the partition function associated with the Lagrangian specified in Eq. (2.1). For this purpose we need to introduce a three-dimensional box and evaluate the thermal trace $Z=\mathrm{Tr}\{\exp(-H/T)\}$, where H is the Hamiltonian on the box. As is well known, the partition function can be represented as a path integral in Euclidean space, the time axis being closed to a circle. In order for this path integral to represent the thermal trace, the gluon field must be periodic up to a gauge transformation, while the quark fields are subject to antiperiodic boundary conditions. It is convenient to choose the same boundary conditions also in the space directions such that the fields then "live" on a four-dimensional torus of size $L_1\times L_2\times L_3\times L_4$. We denote the corresponding volume by $V=L_1L_2L_3L_4$. Setting $a_\mu=n_\mu L_\mu$, where $n_1,\ldots,n_4$ are integers, the boundary conditions then take the form

$$\mathbf{G}_\mu(x+a)=\Omega_a\mathbf{G}_\mu(x)\Omega_a^\dagger-i\Omega_a\partial_\mu\Omega_a^\dagger , \tag{2.3}$$

with $\Omega_a=\Omega_a(x)\in\mathrm{SU}(N_c)$. The corresponding antiperiodicity condition for the quark fields reads

$$q(x+a)=(-1)^{|a|}\Omega_a(x)q(x) , \tag{2.4}$$

with $|a|=n_1+\cdots+n_4$. This condition requires the transition function to obey the composition rule

$$\Omega_{a+b}(x)=\Omega_b(x+a)\Omega_a(x) . \tag{2.5}$$

In geometrical terms we are describing the torus by means of Euclidean space, identifying points if they differ by an element of the lattice $\{a\}$. To describe gauge fields on this geometry, we need to cover Euclidean space with patches, specifying the field on each one of these and providing transition functions which relate the values of the field in the overlap of any two patches. Now, any gauge-field configuration on Euclidean space can be described in terms of a single patch (see Appendix A). For the field at equivalent points to be the same, the values of $\mathbf{G}_\mu(x)$ and $\mathbf{G}_\mu(x+a)$ must then be related by a gauge transformation $\Omega_a(x)$, according to Eq. (2.3). Gauge fields on a torus can therefore be described in terms of a field $\mathbf{G}_\mu(x)$ defined on all of Euclidean space, supplemented by a transition function $\Omega_a(x)$. In this representation the topology of the gauge field exclusively resides in the transition function: If $\mathbf{G}_\mu$ and $\mathbf{G}'_\mu$ obey the periodicity condition (2.3) with the same $\Omega_a(x)$, then the deformation $\xi\mathbf{G}_\mu+(1-\xi)\mathbf{G}'_\mu$ continuously deforms one into the other without changing the transition function. The winding number of the gauge field,

$$\nu=\int_V dx\,\omega(x) , \tag{2.6}$$

is the prototype of a topological invariant. It can explicitly be expressed in terms of the transition function (see Appendix A). The consistency condition (2.5) implies that ν is an integer. In particular, if $\Omega_a(x)=1$, i.e., if $\mathbf{G}_\mu(x)$ is periodic, the winding number vanishes. [When analyzing instanton configurations, it may be more convenient to work in a singular gauge. In this case one is implicitly dealing with more than one patch and the winding number may then show up in the transition functions which connect them rather than in the periodicity condition. Note also that we are considering fermions in the fundamental representation of the color group. The composition rule (2.5) is a condition on the fermion representation of the transition function rather than on the function itself. If the representation of the center of the group is trivial, as it is the case, e.g., for fermions in the adjoint representation, the winding number need not be an integer; cf. Sec. X].

Since the gauge-field configurations occurring in the path integral are characterized by an integer, the partition function involves a sum

$$Z=\sum_{\nu=-\infty}^{\infty} e^{i\nu\theta}Z_\nu . \tag{2.7}$$

The individual terms stem from field configurations of winding number ν:

$$\begin{aligned} Z_\nu&=\int[dG]e^{-S_G}\det(-i\not{D}+\tilde{\mathcal{M}}) ,\\ \tilde{\mathcal{M}}&=\tfrac{1}{2}(1-\gamma_5)\mathcal{M}+\tfrac{1}{2}(1+\gamma_5)\mathcal{M}^\dagger , \end{aligned} \tag{2.8}$$

where S_G is the classical action of the gluon field. Denote the eigenvalues of $\not{D}$ by λ_n. If the winding number is positive (negative), there are $|\nu|$ left-handed (right-handed) zero modes. The nonzero eigenvalues occur in pairs $(\lambda_n,-\lambda_n)$. The determinant is therefore of the form $(\nu>0)$

$$\det(-i\not{D}+\tilde{\mathcal{M}})=(\det{}_f\mathcal{M})^\nu\prod_n{}'\det{}_f(\lambda_n^2+\mathcal{M}\mathcal{M}^\dagger) , \tag{2.9}$$

where $\det_f\mathcal{M}$ stands for the determinant of the $N_f\times N_f$ matrix $\mathcal{M}$ and the product occurring on the RHS only extends over the positive eigenvalues. [If ν is negative, the factor $(\det_f\mathcal{M})^\nu$ is to be replaced by $(\det_f\mathcal{M}^\dagger)^{-\nu}$].

The above representation incorporates a very important property of the partition function which originates in the anomalous Ward identity obeyed by the singlet axial-vector current: Z depends on the quark-mass matrix and vacuum angle only through the product $\mathcal{M}\exp(i\theta/N_f)$—a change in the phase of the quark-mass matrix is equivalent to a change in θ.

The following analysis concerns the properties of the partition function $Z=Z(\theta)$ defined in Eqs. (2.7) and (2.8). In particular, the ground state of our system is the so-called θ vacuum [5]. It is legitimate to consider alternative states. One may, e.g., identify the partition function with the integral $\int_0^{2\pi}d\theta\,Z(\theta)$ or consider states such as $Z(\theta)+Z(\theta+\pi)$. The integral over θ exclusively receives contributions from field configurations with $\nu=0$—we will analyze this sector in detail in Sec. V. The superposition $Z(\theta)+Z(\theta+\pi)$, on the other hand, restricts the path integral to configurations of even winding number. If the quark masses are small, the contributions of the various winding-number sectors to the quark condensate are quite different such that the result strongly depends on the state under consideration. The question of whether the correct "physical" partition function is given by $Z(\theta)$ or by a superposition of the form

$$Z^*=\sum_i c_iZ(\theta_i) \tag{2.10}$$

is, however, an academic one. Since the quarks occuring in nature are massive, different values of θ correspond to different vacuum energies. In the infinite-volume limit, the sum (2.10) is anyway dominated by the term with the smallest vacuum energy and is therefore indistinguishable from the partition function $Z(\theta)$ with the corresponding value of θ. In particular, for real positive quark masses, the states $\int_0^{2\pi} d\theta\, Z(\theta)$, $Z(0)+Z(\pi)$ and $Z(0)$ yield identical results at infinite volume (see also the related discussion in Sec. X). At finite volume and small quark masses where the distinction *is* important, the behavior of the partition function depends in an essential way on the properties of the three-dimensional box used. Since the torus is an academic invention, the finite-volume aspects of the theory cannot be confronted with experiment. They are however of interest, because they shed light on the mathematical structure of the theory. Also, they are relevant for lattice studies which are usually based on a torus. In that framework one may determine the behavior of various observables in the continuum limit for a torus of finite volume. The regularization procedure then unambiguously specifies whether the different winding-number sectors occur with the weight $\exp(i\nu\theta)$, as is the case with $Z(\theta)$, or whether one is dealing with a superposition of the type (2.10).

III. EFFECTIVE THEORY

In the following we exploit the fact that, for large volumes and small quark masses, the properties of the partition function can be analyzed by means of effective-field theory, which describes the low-energy structure of QCD [6–9].

Before embarking on this analysis, a remark concerning the surface effects generated by the walls of the box is in order. If we were only considering bulk properties at infinite volume, then these effects would be entirely irrelevant as they are suppressed by a power of the box size. We are, however, analyzing the properties of the partition function at finite volume, and we therefore need to make sure that finite-size effects are under control. Indeed, for the particular boundary conditions specified in the previous section, this is the case. As shown in Ref. [10], the corresponding effective Lagrangian does not contain surface terms and the effective coupling constants occurring therein are independent of the size of the box. For these boundary conditions, the properties of the partition function at large volumes and small quark masses can therefore be worked out by means of the standard effective chiral Lagrangian. Presumably, our conclusions hold for a more general class of boundary conditions. Note, however, that boundary conditions which introduce a "wall", i.e., violate translation invariance, lead to surface effects which in the presence of long-range forces may modify the structure of the partition function at large volume.

The extension of the torus in the Euclidean time direction determines the temperature, $L_4=1/T$, while the quantities L_1,L_2,L_3 specify the size of the three-dimensional box. We assume that all four sides of the torus are large compared to $1/\Lambda_{\rm QCD}$. The partition function is then dominated by the contributions from the lightest particles of the theory. Their properties strongly depend on the number of flavors and the magnitude of the quark masses. For $N_f \geq 2$, the spontaneous breakdown of chiral symmetry gives rise to N_f^2-1 Goldstone bosons; if the quark masses are small, they represent the lightest particles and hence dominate the behavior of Z at large volume and low temperature. For $N_f=1$, on the other hand, the mass gap is expected to persist in the chiral limit. Presumably, the lightest particle is a pseudoscalar $q\bar{q}$ bound state also in this case, but its mass M_0 is not small in comparison to the scale of the theory. Accordingly, the pressure becomes exponentially small at low temperature, $P \propto \exp(-M_0/T)$, and the finite-size effects generated by the three-dimensional box are also small, of order $\exp(-M_0 L)$. Denoting the energy density of the vacuum by $\epsilon_0=\epsilon_0(me^{i\theta})$ and dropping exponentially small contributions, the partition function becomes

$$Z=\exp\{-V\epsilon_0(me^{i\theta})\} \quad (N_f=1)\ . \tag{3.1}$$

Actually, this representation holds under more general conditions. It is not necessary that the temperature is small compared to the mass gap. If this condition is not met, the partition function, of course, receives contributions from the various excitations of the ground state which depend on the temperature. The formula (3.1), however, still applies, provided that the three-dimensional box is large [11]; the only change brought about by the temperature is that the vacuum-energy density $\epsilon_0(me^{i\theta})$ is to be replaced by the density of the free energy, $f(T,me^{i\theta})$. This has interesting implications for the manner in which temperature affects the distribution of the winding number, but we will not discuss this issue here. In the following we treat the four sides of the torus on equal footing, simply referring to a four-dimensional box of size L.

An analogue of the representation (3.1) also holds if there are two or more flavors. In that case, however, the mass gap is not of the order of the scale of the theory, but is given by the mass of the lighest Goldstone boson and disappears in the chiral limit. We will discuss the modifications required by the occurrence of Goldstone bosons in Secs. VIII and IX. First, we stick to $N_f=1$ and consider the implications of Eq. (3.1) for the distribution of the winding number.

IV. DISTRIBUTION OF THE WINDING NUMBER FOR $N_f=1$

The vacuum energy can be expanded in powers of m, treating the quark-mass term as a perturbation. For $N_f=1$, the spectrum of the theory does not contain massless particles in the chiral limit. The perturbation series in powers of m does therefore not give rise to infrared divergences, such that the expansion of ϵ_0 in powers of m is an ordinary Taylor series:

$$\epsilon_0(me^{i\theta})=\epsilon_{00}-\Sigma\,\mathrm{Re}(me^{i\theta})+O(m^2)\ . \tag{4.1}$$

The first term only affects the overall normalization of the partition function and can be dropped. Disregarding the contributions of $O(m^2)$, the partition function becomes

$$Z=\exp\{\Sigma V\,\mathrm{Re}(me^{i\theta})\}\{1+O(m^2V)\} \ . \qquad (4.2)$$

The low-energy constant Σ is related to the value of the quark condensate in the chiral limit:

$$\langle \bar{q}_L q_R \rangle|_{m=0} = -\frac{1}{V}\frac{\partial}{\partial m^*}\ln Z = -\frac{1}{2}\Sigma e^{-i\theta} \ . \qquad (4.3)$$

The expansion of the scalar and pseudoscalar condensates in powers of m therefore starts with

$$\langle \bar{q}q \rangle = -\Sigma\cos\theta + O(m) \ ,$$
$$\langle \bar{q}i\gamma_5 q \rangle = -\Sigma\sin\theta + O(m) \ . \qquad (4.4)$$

According to Eq. (2.7), the contribution of the sector with winding number ν to the partition function is given by the corresponding coefficient of the Fourier series with respect to θ. Using the explicit representation (4.2), this coefficient is readily worked out, with the result

$$Z_\nu = \left[\frac{m}{|m|}\right]^{\nu} I_\nu(V\Sigma|m|) \ , \qquad (4.5)$$

where $I_\nu(x)=I_{-\nu}(x)$ is Bessel function of imaginary argument:

$$I_\nu(x) = \frac{1}{|\nu|!}\left[\frac{x}{2}\right]^{|\nu|}\left\{1+\frac{x^2}{4(|\nu|+1)}+\cdots\right\} \ , \quad x \gg 1 \ . \qquad (4.6a)$$

$$I_\nu(x) = (2\pi x)^{-1/2}e^{x}\left\{1-\frac{\nu^2-\frac{1}{4}}{2x}+\cdots\right\} \ , \quad x \gg 1 \ . \qquad (4.6b)$$

In particular, in the chiral limit, Z_ν disappears with the power m^ν if ν is positive,

$$Z_\nu = \frac{1}{\nu!}\left[\frac{V\Sigma m}{2}\right]^{\nu}+\cdots \ , \qquad (4.7)$$

and with the power $(m^*)^{|\nu|}$ if ν is negative. This behavior is in accordance with the properties of the Dirac determinant and originates in the fact that on a gluonic background of winding number ν the Dirac operator admits $|\nu|$ zero modes. It may appear like a miracle that the general structure of the partition function at large volume and low temperature, which we are exploiting here, implies the occurrence of zero modes, although it does not explicitly involve the quark and gluon degrees of freedom. The harmony is prestabilized: The representation (4.2) incorporates the anomalous Ward identity obeyed by the axial-vector current. This identity requires both the Fourier coefficient Z_ν and the Dirac determinant associated with a configuration of winding number ν to vanish in the chiral limit, with the same power of the quark mass. The result (4.7) is, however, more specific, as it fixes the coefficient of m^ν in terms of the volume of the box and the vacuum expectation value at $\theta=0$, $\langle 0|\bar{q}q|0\rangle=-\Sigma$. It shows, in particular, that the coefficient grows with the volume: While configurations of winding number $\nu>0$ are suppressed by the power m^ν of the quark mass, they are enhanced by the power V^ν of the volume. If the product $V\Sigma|m|$ is a small number, the mass factor wins and nontrivial topologies become rare, but for $V\Sigma|m|\gg 1$ the volume factor wins such that the suppression is removed.

The angle θ and winding number ν are canonically conjugate quantities. The partition function corresponds to a sharp value of θ; accordingly, the winding number is uncertain. The Fourier decomposition represents the partition function as a coherent sum of terms with a sharp value of ν. In general, the terms occurring in this sum are not real. Note, however, that for $x>0$ the function $I_\nu(x)$ is positive. If the quark mass is real and positive, and if the vacuum angle vanishes (more generally, if $me^{i\theta}$ is positive), the Fourier series therefore admits a probabilistic interpretation: The probability to encounter a field configuration with winding number ν is given by

$$\frac{Z_\nu}{Z} = I_\nu(V\Sigma m)\exp(-V\Sigma m) \ . \qquad (4.8)$$

In the remainder of this section, we discuss the content of this result, setting $\theta=0$ and first taking m real, positive.

The crucial quantity which characterizes the distribution of the winding number is the mean square

$$\langle \nu^2 \rangle = \sum_\nu \nu^2 \frac{Z_\nu}{Z} \ . \qquad (4.9)$$

The Fourier decomposition (2.7) shows that $\langle \nu^2 \rangle$ is the second derivative of the partition function with respect to θ, at $\theta=0$. With the explicit formula (4.2), this gives

$$\langle \nu^2 \rangle = V\Sigma m \ . \qquad (4.10)$$

The mean-square winding number per unit volume, $\langle \nu^2 \rangle / V = \Sigma m$, is referred to as the topological susceptibility. It is independent of the size of the box and disappears if the quark mass tends to zero [12].

In reality, the quark masses are different from zero and the volume of the Universe is pretty large, $V\Sigma m \gg 1$. Let us first determine the shape of the winding-number distribution for this case. Since the mean-square winding number is then large, the probability distribution Z_ν/Z becomes broad, extending to winding numbers of order $|\nu| \sim (V\Sigma m)^{1/2}$. Accordingly, low winding numbers $0, \pm 1, \pm 2, \ldots$ occur with equal probability. The manner in which the distribution falls off at large $|\nu|$ is readily worked out using the uniform asymptotic expansion of Bessel functions [13]:

$$I_\nu(\nu z) \sim \frac{1}{\sqrt{2\pi\nu}}\frac{e^{\nu\eta}}{(1+z^2)^{1/4}}\left\{1+O\left[\frac{1}{\nu}\right]\right\} \ ,$$
$$\eta = (1+z^2)^{1/2} + \ln\frac{z}{1+(1+z^2)^{1/2}} \qquad (4.11)$$

$(\nu>0)$. This representation shows that, in the region $\nu \ll \langle \nu^2 \rangle$, the distribution is Gaussian,

$$\frac{Z_\nu}{Z} = \frac{1}{(2\pi\langle \nu^2 \rangle)^{1/2}} e^{-\nu^2/2\langle \nu^2 \rangle} \ , \quad V\Sigma m \gg 1 \ , \qquad (4.12)$$

while winding numbers comparable to or larger than $\langle \nu^2 \rangle$ are very improbable. As a check, one may verify that the distribution (4.12) indeed sums up to 1 and yields the proper mean-square deviation.

The above result for the behavior of the distribution in the limit $V \gg 1/\Sigma m$ is intuitively reasonable. Dividing space into a large number of cells and assuming that the field fluctuations within two different cells are independent from one another, one expects that the square of the total winding number is proportional to the number of cells, i.e., to the volume. Moreover, the central limit theorem of probability theory predicts that, irrespective of the distributions in the various cells, the fluctuations in the total winding number are Gaussian.

In the opposite limit $\langle \nu^2 \rangle = V\Sigma m \ll 1$, in particular for massless quarks at finite volume, the winding-number distribution is concentrated at $\nu=0$: $Z_0 \simeq 1, Z_{\pm 1} \simeq V\Sigma m/2, \ldots$. The mean value of ν^2 stems from configurations with $\nu = \pm 1$, and the same two sectors also dominate the quark condensate. Clearly, for $V\Sigma m \ll 1$, the winding number plays a crucial role.

V. TOPOLOGICALLY TRIVIAL CONFIGURATIONS

Suppose that the sum over all gluon-field configurations is restricted to those with $\nu=0$. The partition function is then given by $Z_0 = I_0(V\Sigma m)$, and the quark condensate becomes

$$\langle \bar{q}q \rangle_{\nu=0} = -\Sigma \frac{I_0'(V\Sigma m)}{I_0(V\Sigma m)} \ . \tag{5.1}$$

Since $I_0(x)$ is even in x, this expression tends to zero if $m \to 0$, in contrast with the condensate of the full theory. In the region $V\Sigma m \ll 1$, the restriction to the sector $\nu = 0$ therefore changes the properties of the theory in an essential way.

These findings are paradoxical. The winding number is a global concept, and one should expect that, since the theory has a mass gap M_0, the value of the condensate at a given point should only be sensitive to the behavior of the gluon fields within a ball of size $1/M_0$ around this point. It should not be affected by the presence of a few instantons behind the moon and should therefore be the same whether or not one restricts the winding number of the Universe to zero. The asymptotic representation of the Bessel function shows that the quark condensate does have this property if the volume is large, but the asymptotic behavior only sets in for $V \gg 1/\Sigma m$—it is not the mass gap which sets the scale here. The resolution of the puzzle is the following. Long-range correlations are absent in the full theory, but they do occur if the integration over the gluon field is restricted to a fixed winding number. In fact, the long-range correlations are generated precisely by the constraint $\int_V dx\, \omega = \nu$. To see this, consider the two-point function of the operator $\bar{q}q$ in the sector with winding number zero:

$$\sigma_0(x-y) = \langle \bar{q}(x)q(x)\bar{q}(y)q(y) \rangle_{\nu=0} \ . \tag{5.2}$$

The space integral of this quantity represents the second derivative of the partition function with respect to the quark mass:

$$\int_V dx\, dy\, \sigma_0(x-y) = \frac{1}{Z_0} \frac{\partial^2 Z_0}{(\partial m)^2} \ . \tag{5.3}$$

Inserting the explicit expression for Z_0 and taking the limit $m \to 0$, this relation implies that the volume integral of the correlation function is proportional to the volume:

$$\int_V dx\, \sigma_0(x) = \tfrac{1}{2}\Sigma^2 V \ . \tag{5.4}$$

Hence $\sigma_0(x)$ does not fall off within a distance of order $1/M_0$—if it did, the space integral would become volume independent for $V \gg M_0^{-4}$. The result (5.4) implies that, for distances large compared to $1/M_0$, the correlator $\sigma_0(x)$ tends to $\frac{1}{2}\Sigma^2$. Note that the nonzero asymptotic value is not due to a disconnected part: Since we are restricting the gluon-field configurations to winding number zero, the expectation value of $\bar{q}q$ vanishes if the quarks are massless [see Eq. (5.1)]. This should be confronted with the situation in the full theory, where the condensate does not disappear in the chiral limit, because the sectors $\nu = \pm 1$ generate nonzero contributions, while the correlator picks up terms both from $\nu=0$ and ± 2. In the full theory, the correlator does obey the cluster-decomposition property, tending to the square of the condensate if $x \to \infty$. (This is exactly what happens in the two-dimensional Schwinger model—for a thorough discussion of the fermion correlators in this model at finite volume, see Ref. [14].)

We conclude that if the integral over the gluon field is restricted to configurations with trivial topology, $\nu=0$, one obtains a different theory, involving long-range correlations. The corresponding partition function $Z_0 = I_0(V\Sigma m)$ is reminiscent of models which undergo spontaneous symmetry breakdown. Two crucial features are in common. First, models with spontaneous symmetry breaking also involve long-range correlations due to the presence of massless particles, Goldstone bosons. Second, the quark condensate (5.1) behaves in the same way as the order parameter in such models. The sign of $\langle \bar{q}q \rangle_{\nu=0}$ changes when the quark mass changes sign (the same occurs with the spontaneous magnetization of a ferromagnet under a sign change of the external magnetic field). In the chiral limit at finite volume, $\langle \bar{q}q \rangle_{\nu=0}$ disappears (which is reminiscent of the absence of spontaneous symmetry breaking at finite volume). Indeed, there are local-field theories for which the partition function at large volume is given precisely by the Bessel function $I_0(V\Sigma m)$. This behavior results under the following circumstances. Consider a field theory with U(1) symmetry and assume that it contains an operator $O(x)$ which transforms according to the representation $O(x) \to e^{i\alpha}O(x)$ of U(1). Suppose that the ground state spontaneously breaks U(1) symmetry with a nonzero value of the order parameter $\Sigma = |\langle 0|O(x)|0\rangle|$.[2] Denote the

[2]The simplest example of such a theory is the textbook model of a complex scalar field:

$$\mathcal{L} = \partial_\mu \varphi\, \partial_\mu \varphi^* - \lambda(\varphi\varphi^* - \Sigma^2)^2 \ , \tag{5.5}$$

with $O(x) = \varphi(x)$.

U(1)-invariant Lagrangian by $\mathcal{L}_0$ and introduce a symmetry-breaking term proportional to $O(x)$:

$$\mathcal{L}=\mathcal{L}_0+\tfrac{1}{2}\{m^*O(x)+mO^\dagger(x)\} \ . \tag{5.6}$$

The spontaneous breakdown of the symmetry gives rise to Goldstone bosons which dominate the partition function at large volume. For small m and large V, the partition function can explicitly be worked out by means of the corresponding effective Lagrangian which involves a single Goldstone field. One indeed finds that the leading term is given by the Bessel function $I_0(V\Sigma m)$. Does the restriction of the gluon field to configurations with trivial topology turn QCD into such a theory? If so, we would be dealing with two different realizations of QCD with $N_f=1$: (i) The gluon field fluctuates with arbitrary topology, and U(1) symmetry is ruined by the axial anomaly which ensures that the η' is massive; (ii) only trivial topologies occur such that the anomaly is prevented from having physical effects, U(1) symmetry being broken only spontaneously, and the spectrum contains a Goldstone boson, i.e., a massless η'.

The picture may be suggestive, but it is not correct. The long-range correlations which manifest themselves in Z_0 do not originate in Goldstone bosons. Otherwise, the partition function would necessarily also contain contributions from particles moving at small, but nonzero momenta of order $p\sim 1/L$, where L is the size of the box. The perturbation series associated with the effective Lagrangian amounts to an expansion in powers of $1/L^2$. The leading term in this expansion exclusively involves Goldstone bosons with $p=0$, but the occurrence of modes with nonzero momenta reveals itself a first nonleading order, through a one-loop graph which generates a power correction of order $1/L^2$. The point here is that the large volume expansion of Z_0 does not contain such power corrections. The long-range correlations seen in Z_0 are those of a theory which contains a single degree of freedom with $p=0$ rather than the full spectrum of a massless particle.

In fact, the origin of the zero mode is easily identified. The inversion of the Fourier decomposition,

$$Z_0=\frac{1}{2\pi}\int_0^{2\pi}d\theta\, Z(\theta) \ , \tag{5.7}$$

shows that the projection onto field configurations with $\nu=0$ is given by the integral of the full partition function over the vacuum angle. Hence the long-range correlations originate in the integral over θ. Indeed, the fact that the vacuum angle in some respects behaves like the order parameter of a spontaneously broken symmetry was noted in the literature [15]. The operator which generates a shift in the vacuum angle is referred to as $Q^5_{\rm sym}$ and differs from the axial U(1) charge of the quarks by a Chern-Simons term which stems from the anomaly. For massless quarks, $Q^5_{\rm sym}$ is conserved. In this language the θ vacuum may be interpreted as a state which spontaneously breaks the symmetry generated by $Q^5_{\rm sym}$ and one may then wonder why this symmetry breakdown does not give rise to a Goldstone boson. The reason is the following. The Goldstone degrees of freedom are related to symmetry operations for which the parameters of the Lie group are allowed to slowly vary in space. The corresponding generator is given by an integral over the charge density weighted with a space-dependent factor. In other words, the Goldstone theorem requires the occurrence of a conserved current, not only of a conserved charge. The Chern-Simons "charge" can be written as an integral over a local current, but as is well known, this current is not gauge invariant; i.e., it generates unphysical states. One may allow the parameter θ occurring in the Lagrangian (2.1) to depend on space, but the θ term then fails to be a total derivative—there is no way to make sense out of $Q^5_{\rm sym}$ in terms of an integral over a local charge density. There is, therefore, no reason for the theory to contain massless particles with the quantum numbers of axial U(1): The anomaly does take care of the U(1) problem.

The situation is rather similar to the one in general relativity, where energy-momentum conservation cannot be formulated in terms of a sensible local conservation law. Although the total energy and momentum of the system are conserved if the gravitational field vanishes at spatial infinity, they cannot be localized—what is conserved [16] is not a tensor.

Finally, we emphasize that the long-range correlations mentioned above only occur if the parameter $V\Sigma m$ is small. If the quark mass is given some small nonzero value, the correlations disappear as the volume tends to infinity. Indeed, the above formulas show that for the density of free energy, $f=-V^{-1}\ln Z$, it is irrelevant whether one sums over all gluon-field configurations with $\theta=0$ or restricts oneself to the topologically trivial ones, provided $V\Sigma m \gg 1$. Up to corrections of order $1/V$, the free-energy density is given by $-\Sigma m$, in either case. Unless the quark mass is strictly zero, the winding number becomes irrelevant if the volume is large enough.

VI. SPECTRUM OF THE DIRAC OPERATOR FOR $N_f=1$

The long-range correlations discussed in the preceding section are closely related to the fact that the Dirac operator admits small eigenvalues. In the present section, we explore this relation in detail, using the explicit results for the partition function to obtain information about the eigenvalue distribution. Throughout the following we consider a given winding-number sector, i.e., look at the quantity Z_ν, for some fixed value of ν which we, for convenience, assume to be positive or zero. Since the phase of the quark mass then only enters in a trivial way, we take m real and positive.

The quantity Z_ν is the functional integral

$$Z_\nu=m^\nu\int_\nu [dG]e^{-S_G}\prod_n{}'(\lambda_n^2+m^2) \ , \tag{6.1}$$

where the product only extends over the positive eigenvalues of the Dirac operator and the integral is taken over gluon-field configurations with a given winding number ν. The formula shows that, except for the factor m^ν, the dependence of Z_ν on the quark mass arises from an average of the factor $\prod_n'(1+m^2/\lambda_n^2)$ over all gluon-field configurations with winding number ν. It is convenient to denote this average by $\langle\langle \cdots \rangle\rangle_\nu$:

$$\langle\langle f \rangle\rangle_\nu \equiv \frac{\int [dG] e^{-S_G} \prod_n' \lambda_n^2 f}{\int [dG] e^{-S_G} \prod_n' \lambda_n^2} . \tag{6.2}$$

In this notation, Eq. (6.1) amounts to the exact relation

$$\left\langle\left\langle \prod_n' \left[1 + \frac{m^2}{\lambda_n^2} \right] \right\rangle\right\rangle_\nu = \frac{m^{-\nu} Z_\nu(m)}{\lim_{m\to 0}[m^{-\nu} Z_\nu(m)]} . \tag{6.3}$$

Inserting the explicit representation (4.5) for Z_ν, which applies if the box is large compared to the scale of the theory, we obtain

$$\left\langle\left\langle \prod_n' \left[1 + \frac{m^2}{\lambda_n^2} \right] \right\rangle\right\rangle_\nu = \nu! \left[\frac{2}{x} \right]^\nu I_\nu(x) , \tag{6.4}$$

with $x \equiv V\Sigma m$.

The qualitative content of this result is best seen by invoking the infinite-product representation of the Bessel function in terms of its zeros. Denoting the positive zeros of the function $J_\nu(x)$ by $\xi_{n,\nu}$,

$$J_\nu(\xi_{n,\nu}) = 0, \quad n = 1, 2, \ldots , \tag{6.5}$$

the function $I_\nu(x)$ can be written as

$$I_\nu(x) = \frac{1}{\nu!} \left[\frac{x}{2} \right]^\nu \prod_{n=1}^\infty \left[1 + \left[\frac{x}{\xi_{n,\nu}} \right]^2 \right] . \tag{6.6}$$

The relation (6.4) then takes the form

$$\left\langle\left\langle \prod_n' \left[1 + \frac{m^2}{\lambda_n^2} \right] \right\rangle\right\rangle_\nu = \prod_{n=1}^\infty \left[1 + \left[\frac{V\Sigma m}{\xi_{n,\nu}} \right]^2 \right] . \tag{6.7}$$

This shows that the result of the path integral over the gluon field is the same as if there was a single field configuration with eigenvalues distributed according to[3]

$$\bar{\lambda}_n = \pm \frac{\xi_{n,\nu}}{V\Sigma} . \tag{6.8}$$

As anticipated in Sec. I, the eigenvalues are inversely proportional to the volume. Moreover, the asymptotic formula

$$\xi_{n,\nu} \simeq \pi(n + \tfrac{1}{2}\nu - \tfrac{1}{4}) \tag{6.9}$$

shows that for eigenvalues which are large compared to $1/V\Sigma$, the spectrum is equidistant with $\Delta\lambda = \pi/V\Sigma$. The mean level density per unit volume is therefore independent of the winding number and indeed approaches the value

$$\rho(0) = \Sigma/\pi = |\langle 0|\bar{q}q|0\rangle|/\pi$$

in the infinite-volume limit. The distribution of the lowest few eigenvalues, however, is sensitive to the winding number. In particular, there are ν zero modes. As a compensation, the spectrum is diluted in the vicinity of $\lambda = 0$, the zeros of the Bessel function being shifted toward larger values if ν is increased.

In reality, the eigenvalues of the Dirac operator, of course, depend on the gluon field under consideration. The explicit formula (6.8) merely indicates that the spectrum fluctuates around these values. On a quantitative level, the content of Eq. (6.4) can be expressed in terms of inverse moments of the eigenvalue spectrum. Expanding both sides of this equation in a Taylor series in powers of m^2 and comparing coefficients, we obtain a string of sum rules involving inverse powers of λ_n. Expressed in terms of the corresponding dimensionless eigenvalues $l_n \equiv \lambda_n V\Sigma$, we get

$$\left\langle\left\langle \sum_n' \frac{1}{l_n^2} \right\rangle\right\rangle_\nu = \frac{1}{4(\nu+1)} , \tag{6.10a}$$

$$\left\langle\left\langle \left[\sum_n' \frac{1}{l_n^2} \right]^2 - \sum_n' \frac{1}{l_n^4} \right\rangle\right\rangle_\nu = \frac{1}{16(\nu+1)(\nu+2)} , \tag{6.10b}$$

etc., where the sums only extend over positive eigenvalues. The explicit formula (6.8), according to which the dimensionless mean eigenvalues are given by the zeros of the Bessel function, $\bar{l}_n = \pm\xi_{n,\nu}$, holds in the sense of these inverse moments. In particular, the mean value of $\sum_n' 1/l_n^2$ is given by $\sum_n 1/\xi_{n,\nu}^2 = 1/4(\nu+1)$.

The results (6.10) are nontrivial and can, in principle, be checked in lattice experiments. First lattice measurements of the spectrum of the Dirac operator for some field-theory models have already appeared [17]. The recent progress made in calculations with dynamical fermions [18] should allow one to obtain similar eigenvalue distributions also for QCD and to perform an experimental check of the sum rules (6.10). The same applies to the topological susceptibility (4.10). For lattice measurements of $\langle \nu^2 \rangle$ in pure Yang-Mills theory, see [19].

A remark concerning ultraviolet divergences is in order here. The above discussion ignores the fact that the product $\prod_n'(1 + m^2/\lambda_n^2)$ requires regularization, and the same applies to sums such as $\sum_n' 1/\lambda_n^2$. How can we claim that, in the infinite-volume limit, the results (6.10) are exact if the LHS involves divergent sums, while the RHS is finite? The point is that the divergences only affect the corrections of order $1/V$. To verify this statement, we first note that our analysis only concerns those eigenvalues of the Dirac operator which are small compared to the scale of the theory. Our neglect of the higher-order terms in Eq. (4.2) is justified only if the quark mass is small, $m \ll \Lambda_{QCD}$. This restriction prevents us from analyzing the spectrum in the region $\lambda \gtrsim \Lambda_{QCD}$ responsible for the occurrence of divergences. For a given field configuration, the level density at $|\lambda| \to \infty$ is given by phase space. The number of levels contained in the interval $(\lambda, \lambda + \Delta\lambda)$ tends to

$$\Delta n = \frac{N_c}{4\pi^2} V |\lambda|^3 \Delta\lambda , \tag{6.11}$$

where N_c is the number of colors. Accordingly, the sum $\sum_n' 1/\lambda_n^2$ contains a quadratic divergence $\propto \Lambda_0^2 V$. As discussed in Sec. I, the same divergence also shows up in the

[3]For similar considerations in the context of noncompact QED on a lattice, see Göckeler *et al.*, in Ref. [17].

vacuum energy and is removed by adjusting the cosmological constant. The contributions from small eigenvalues, on the other hand, grow with the square of the volume. If the volume is large, these contributions therefore dominate the sum, large eigenvalues only generating corrections of relative order $1/V$. Thus, in order to compare our predictions (6.10) with experiment, one should either make sure that the condition $Va^2\Sigma^2 \gg 1$ is satisfied (a is the lattice spacing) or restrict the sum to small eigenvalues, $\lambda_n \lesssim \Lambda_{QCD}$.

VII. QUARK CONDENSATE AND SIGN OF THE QUARK MASS

Let us now discuss the mechanism responsible for the formation of a quark condensate in more detail. For $\theta=0$ and a real, positive quark mass, the decomposition of the partition function into contributions from the various topological sectors becomes

$$Z=I_0(V\Sigma m)+2\sum_{\nu>0} I_\nu(V\Sigma m) . \tag{7.1}$$

The derivative of $I_\nu(V\Sigma m)$ with respect to m contains two terms:

$$I'_\nu(x)=\frac{\nu}{x}I_\nu(x)+I_{\nu+1}(x) . \tag{7.2}$$

The infinite-product representation (6.6) shows that the first term stems from the zero-mode factor x^ν, while the term $I_{\nu+1}(x)$ comes from the derivative of the remainder and is related to the nonzero modes. Thus the relative fraction of the zero-mode contribution to the quark condensate is given by

$$\begin{aligned} f(x)=\frac{\langle \bar{q}q \rangle_{\text{zero modes}}}{\langle \bar{q}q \rangle}&=2e^{-x}\sum_{\nu>0}\frac{\nu I_\nu(x)}{x} \\ &=e^{-x}\{I_0(x)+I_1(x)\} . \end{aligned} \tag{7.3}$$

At small values of $x \equiv V\Sigma m$, the contribution of the zero modes dominates, $f(x)=1-\frac{1}{2}x+O(x^2)$. In this case the leading term arises from $\nu=\pm 1$, in agreement with the standard instanton picture.[4] In the limit $x \gg 1$, however, the zero-mode fraction is negligible, $f(x)\simeq\sqrt{2/\pi x}$. The condensate is determined mainly by the nonzero modes, and the contributions from all winding numbers in the range $|\nu| \lesssim \sqrt{x}$ are equally important.

[4]At small temperatures and large volumes, the quasiclassical approximation underlying the instanton picture of QCD is not justified—the integrals acquire their main contributions from instantons of large size where the quasiclassical parameter $8\pi^2/g^2(\rho)$ fails to be large (at high temperatures, the picture may be adequate, particularly for $N_f=1$, where the value for the quark condensate does not depend on the order in which the limits $m\rightarrow 0$ and $V\rightarrow\infty$ are performed [20]). Our own treatment is in no way based on the quasiclassical approximation. The relevant gluon-field configurations may be quite different from instantons. What counts here is that configurations of winding number $\nu=\pm 1$ carry one quark zero mode. In the small-x limit, the quark condensate exclusively stems from this particular mode.

An important remark concerning the sign of the quark mass is in order here. The above simple physical picture only holds if the Fourier decomposition of the partition function admits a probability interpretation, i.e., if $me^{i\theta}$ is real and positive. Otherwise, it is not meaningful to speak of a probability distribution of gluon-field configurations. In fact, arguments based on this concept can even be misleading. As an illustration, consider the large volume relation (1.5) which expresses the condensate in terms of the mean level density. Let us fix the vacuum angle at $\theta=0$, and let us see what happens if the quark mass is allowed to take negative values. Equation (1.5) indicates that the sign of the condensate is opposite to the sign of m. If this relation holds, the condensate must therefore flip sign as m passes through zero. This behavior, however, is in contradiction with the general properties of the partition function discussed in Sec. IV. According to that analysis, the value of the condensate in the chiral limit is given by $\langle \bar{q}q \rangle = -\Sigma\cos\theta$, irrespective of whether the limit is taken from positive or negative values of m. The contradiction originates in the notion of a mean level density $\rho(\lambda)$, which implicitly requires an average over the various gluon-field configurations, weighted with the appropriate probability. If the mass is negative, the decomposition $Z=\sum_\nu Z_\nu$ becomes an alternating series, odd winding numbers generating negative contributions. In this case the sum over all field configurations thus contains negative terms and probabilistic notions such as the mean level density lose their meaning. In particular, the large volume relation (1.5) only holds if $me^{i\theta}$ is positive. The same remark also applies to the fractions of the quark condensate due to zero and nonzero modes. If the mass is negative and large compared to $1/V\Sigma$, Eq. (7.3) shows that the zero-mode "fraction" is not suppressed at all, but instead becomes exponentially large. It is canceled out almost completely by an exponentially large negative "fraction" due to nonzero modes—by definition, the two contributions add up to 1.

At first sight this suggests that, in QCD with one quark flavor, a change in the sign of the quark mass drastically changes the physical content of the theory. The phenomenon originates in the anomaly of the axial-vector current. The corresponding Ward identity implies that the partition function depends on m and θ only through the combination $me^{i\theta}$. Hence a change in the sign of the quark mass is equivalent to going from $\theta=0$ to π and therefore amounts to a change of sign of the contributions associated with an odd winding number. If only topologically trivial configurations were relevant, the partition function would not discriminate between $\pm m$ and the sign of the quark condensate would indeed be opposite to the sign of m [see Eq. (4.13)]. The full theory exhibits a different behavior because nontrivial topologies cannot be ignored.

Note, however, that these conclusions only hold for the specific state defined in Eqs. (2.7) and (2.8). As mentioned in Sec. II, one may envisage alternative states. Consider, e.g., the two inequivalent theories characterized by (θ,m) and $(\theta,-m)$, respectively. Each of these theories acts in its own space of states. Take the direct sum of the two spaces. The corresponding partition function is given by

$$Z_{\rm tot}(\theta,m)=Z(\theta,m)+Z(\theta,-m)\ . \tag{7.4}$$

The point is that, if the volume is sent to infinity, only one of the two terms survives—the one with the lower vacuum energy. In particular, for $\theta=0$, the large volume limit of $Z_{\rm tot}$ coincides with $Z(0,m)$ if $m>0$ and with $Z(0,-m)$ if $m<0$. The construction thus allows the system to choose between two different ground states, one for which $\langle\bar{q}q\rangle=-\Sigma$ and one where $\langle\bar{q}q\rangle=+\Sigma$. The system picks the first state if m is positive and the second one for $m<0$. In view of the identity $Z(\theta,-m)=Z(\theta+\pi,m)$, the quantity $Z_{\rm tot}$ is periodic in θ with period π. Its Fourier decomposition therefore only contains even winding numbers,

$$Z_{\rm tot}=2\sum_{\nu=2k} Z_\nu e^{i\nu\theta}\ , \tag{7.5}$$

a further illustration of the statement that the winding number is an important quantum number only at finite volume. The example demonstrates that the question of what happens if the sign of the quark mass is flipped can unambiguously be answered only at finite volume where $Z_{\rm tot}$ clearly differs from either one of the two irreducible components of which it is composed.

Incidentally, the issue also arises in QED. The corresponding partition function on a torus involves a sum over topologically inequivalent configurations of the electromagnetic field and sectors with an odd winding number again generate negative contributions if the electron mass is negative. At finite volume the two theories with $\pm m_e$ are therefore inequivalent.

VIII. SEVERAL FLAVORS

Let us now consider real QCD, which involves several light-quark flavors. There is a very substantial difference between this theory and the single-flavor case discussed above. For $N_f=1$, the formation of a quark condensate does not violate any symmetry. The U(1) axial symmetry is broken explicitly by the quantum anomaly. The mass of the pseudoscalar meson persists in the chiral limit. The spectrum involves a gap, and the partition function obeys the large volume representation (3.1).

If $N_f\geq 2$, only the divergence of the *singlet* axial-vector current contains an anomaly. In the chiral limit, the group $\mathrm{SU}_R(N_f)\times\mathrm{SU}_L(N_f)\times\mathrm{U}_V(1)$ remains a true symmetry of the Lagrangian. The formation of a quark condensate breaks this group down to $\mathrm{U}_V(N_f)$. As a consequence, the spectrum of the theory now involves N_f^2-1 massless Goldstone bosons. Away from the chiral limit, these particles do carry mass, but if the quarks are light, the mass is small. In a basis where the quark-mass matrix is diagonal with positive eigenvalues $m_u, m_d, \ldots$, the mass of the Goldstone mode with the quantum numbers of $\bar{d}u$, e.g., is proportional to $(m_u+m_d)^{1/2}$. For $\theta=0$, the mass is given by

$$M_{\pi^+}^2=(m_u+m_d)\frac{\Sigma}{F^2}+O(m^2)\ , \tag{8.1}$$

where Σ is the quark condensate in the chiral limit at infinite volume,

$$\langle 0|\bar{u}u|0\rangle=\langle 0|\bar{d}d|0\rangle=\cdots=-\Sigma\ , \tag{8.2}$$

and $F\simeq 93$ MeV is the pion-decay constant. If the box is taken large compared to the Compton wavelength of the lightest Goldstone boson, $M_\pi L\gg 1$, then the finite-size effects are exponentially small, of order $\exp(-M_\pi L)$. In this case the partition function is again given by Eq. (3.1). The results established in the preceding sections for $N_f=1$, however, rely to a considerable extent on the behavior of the partition function in the region $V\Sigma m\lesssim 1$. There, the product $M_\pi^2L^2\sim V\Sigma m/F^2L^2$ is a small number, because we are interested in boxes which are large compared to the scale of the theory, $FL\gg 1$. In the region $V\Sigma m\lesssim 1$, the finite-size effects are therefore not suppressed, the logarithm of the partition function is not extensive, and the representation (3.1) does not apply. We need to account explicitly for the Goldstone bosons which are responsible for the fact that the mass gap disappears in the chiral limit. This was done in Refs. [7, 10] where it was shown that, for large volumes and small quark masses, the QCD partition function coincides with the partition function of the standard effective low-energy theory. In that framework, the Goldstone bosons are described by a matrix field $U(x)\in\mathrm{SU}(N_f)$, and the partition function is given by a path integral over this field:

$$Z=\int[dU]\exp\left\{-\int_V d^4x\,\mathcal{L}_{\rm eff}(U,\partial U,\partial^2 U,\ldots;\mathcal{M}e^{i\theta/N_f})\right\}\ . \tag{8.3}$$

At leading order in an expansion in powers of derivatives and powers of the quark-mass matrix $\mathcal{M}$, the effective Lagrangian is of the form

$$\mathcal{L}_{\rm eff}=\frac{F^2}{4}\mathrm{tr}(\partial_\mu U^\dagger\partial_\mu U)-\Sigma\,\mathrm{Re}\{e^{i\theta/N_f}\mathrm{tr}(\mathcal{M}U^\dagger)\}+\cdots\ . \tag{8.4}$$

As far as the static or quasistatic properties are concerned, the effective theory is a full substitute for QCD. In particular, for large volumes and small quark masses, the properties of the partition function are determined by the two effective coupling constants F and Σ, which characterize the leading term in the derivative expansion of the effective Lagrangian—higher-derivative terms only generate corrections involving powers of $1/L^2$ or powers of $\mathcal{M}$. Note that the effective coupling constants are independent of the box size. In the representation (8.3), the integral extends over all pion fields which are periodic in the four directions of Euclidean space—the size of the box only enters through this boundary condition.

For $\theta=0$ and a diagonal, positive mass matrix, the minimum of the action occurs at $U(x)=1$. If the product $V\Sigma\mathcal{M}$ is large, the minimum is narrow and the path integral can be evaluated by expanding the field

$$U(x)=e^{i\pi(x)/F} \tag{8.5}$$

in powers of the traceless Hermitian matrix $\pi(x)$. In the action the terms quadratic in $\pi(x)$ describe a set of N_f^2-1 free fields, while the higher-order terms represent interaction vertices. The path integral then boils down to a Gaussian integral over a polynomial of the field $\pi(x)$, which can be worked out in the standard fashion. The only modification brought about by the box is that the standard meson propagators are replaced by the corresponding Green's functions on the torus, e.g.,

$$G_\pi(x)=\frac{1}{V}\sum_p \frac{e^{ipx}}{M_\pi^2+p^2}\ , \tag{8.6}$$

where the components of the vector p_μ are integer multiples of $2\pi/L_\mu$ and the meson masses are the same as at infinite volume.

The standard chiral perturbation series is ordered according to powers of momenta, the meson masses $M_{\pi^+}, M_{\pi^0}, \ldots$ and the inverse box sides $1/L_\mu$ counting as quantities of $O(p)$. The leading contribution to the partition function stems from the classical action. The fluctuations of the meson field only show up at first nonleading order, through a one-loop graph given by the determinant of the differential operator $D=-\partial^2+M^2$, which characterizes the kinetic term in the effective Lagrangian (M is the meson-mass matrix):

$$Z=\exp\{-V\mathcal{L}_{\text{eff}}|_{U=1}\}(\det D)^{-1/2}\{1+O(p^2)\}\ . \tag{8.7}$$

The classical action (tree graph) is of $O(p^{-2})$, and the determinant (one loop) is of order 1, while graphs with two or more loops only generate corrections of $O(p^2)$. The one-loop divergences contained in $\det D$ require regularization. Since the divergences are the same as at infinite volume, they can be absorbed in the cosmological constant contained in $\mathcal{L}_{\text{eff}}$.

For a general mass matrix and a nonzero vacuum angle, the classical solution does not sit at $U=1$. The value of the effective Lagrangian at the minimum determines the vacuum-energy density $\epsilon_0=\epsilon_0(\mathcal{M}e^{i\theta/N_f})$, and the curvature at the minimum fixes the Goldstone masses $M_i=M_i(\mathcal{M}e^{i\theta/N_f})$. In terms of these quantities, the chiral representation for the partition function becomes [7]

$$Z=\exp\left\{-V\epsilon_0+\tfrac{1}{2}V\sum_{i=1}^{N_f^2-1} g_0(M_i,L_1,\ldots,L_4)\right\}. \tag{8.8}$$

The quantity $-\frac{1}{2}g_0$ is the free-energy density of a Bose gas of free particles on a torus, and the sum extends over the N_f^2-1 Goldstone flavors. The shape of the box only enters through the kinematical function g_0:

$$g_0(M,L_1,\ldots,L_4)=\frac{1}{16\pi^2}\int_0^\infty \frac{dt}{t^3}e^{-tM^2}\sum_{a\neq 0}e^{-a^2/4t}\ , \tag{8.9}$$

where the components of the vector a_μ run over integer multiples of the box sides L_μ. If the box is large compared to the meson Compton wavelength, $ML\gg 1$, the function g_0 becomes exponentially small and the partition function therefore takes the form $Z=\exp(-V\epsilon_0)$, where ϵ_0 is independent of the box. This verifies that, for nonzero quark masses, the Goldstone excitations freeze—the thermal trace $Z=\text{Tr}\{\exp(-\beta H)\}$ reduces to the contribution from the ground state. In the framework of the effective theory, the vacuum-energy density ϵ_0 receives contributions both from the classical effective Lagrangian and loop graphs. Normalizing the cosmological constant such that the vacuum energy vanishes in the chiral limit and setting $\theta=0$, the expansion of ϵ_0 in powers of $\mathcal{M}$ starts with $\epsilon_0=-\Sigma\,\text{tr}\mathcal{M}+O(\mathcal{M}^2)$. Hence, if the box becomes large, $ML\gg 1$, the quark condensate approaches the infinite-volume result (8.2), up to symmetry-breaking effects of order $\mathcal{M}$.

If the product $V\Sigma\mathcal{M}$ is not large, the path integral is not squeezed to the vicinity of $U(x)=1$. In particular, if $\mathcal{M}$ vanishes, the different directions in group space become equally likely—the symmetry is restored. In the standard chiral perturbation series, the phenomenon manifests itself through the fact that the loop expansion develops infrared singularities if the pion mass tends to zero. Indeed, the formula (8.6) shows that the contribution to the propagator from the mode with $p=0$ is given by $1/M_\pi^2V$ and therefore blows up for $M_\pi\to 0$. Accordingly, graphs where mesons with $p=0$ propagate around a loop diverge in the chiral limit. The disease is of technical nature and is easily cured. It originates in the assumption that the path integral (8.3) can be evaluated by expanding the integrand around the Gaussian term which stems from the quadratic part of the effective Lagrangian. Since the mode with $p=0$ does not carry kinetic energy if the pion mass is turned off, the fluctuations in this mode are not Gaussian. The problem only concerns the zero mode; the amplitudes of the modes with $p\neq 0$ do show up in the quadratic part of the action even if the pion mass vanishes. There is a standard technique which allows one to deal with problems of this type: One treats the zero-mode amplitudes as collective variables, setting $U(x)=U_0U_1(x)$, where U_0 is the space-independent field associated with the zero-mode fluctuations, while $U_1(x)$ describes the modes with $p\neq 0$. The latter do not acquire large amplitudes and can therefore be analyzed perturbatively. Ignoring the fluctuations in the nonzero modes altogether, the path integral reduces to an ordinary integral over the collective variable U_0 [7,9,21]:

$$Z = A \int_{SU(N_f)} d\mu(U_0) \exp[V\Sigma \, \mathrm{Re}\{e^{i\theta/N_f} \mathrm{tr}(\mathcal{M} U_0^\dagger)\}] \; . \tag{8.10}$$

The integral extends over the group $SU(N_f)$, and $d\mu(U_0)$ is the corresponding Haar measure. The nonzero-mode fluctuations and higher-order terms in the derivative expansion of the effective Lagrangian generate corrections to this result. As is the case with the representation (8.8), the formula (8.10) becomes exact in the limit $V \to \infty$. The difference between Eqs. (8.8) and (8.10) originates in the fact that we are dealing with a double series, involving powers of $\mathcal{M}$ and of $1/L^2$. In the standard chiral expansion (8.8), $\mathcal{M}$ and $1/L^2$ are treated as quantities of the same order, such that the product $\mathcal{M}L^2 \propto (ML)^2$ counts as a term of order L^0; the series is of the form

$$\ln Z = L^2 f_{-1} + f_0 + L^{-2} f_1 + \cdots \; ,$$

where the coefficients f_n are functions of $\mathcal{M}L^2$. The transformation of the path integral which leads to Eq. (8.10) reorders this series: $\mathcal{M}$ is now treated as a term of order L^{-4}, such that the variable $V\Sigma\mathcal{M}$ counts as a quantity of order 1. The reordered series is of the form $\ln Z = \tilde{f}_0 + L^{-2}\tilde{f}_1 + \cdots$, where the coefficients $\tilde{f}_n$ are functions of $\mathcal{M}L^4$. Equation (8.10) specifies the leading term $\tilde{f}_0$.

As shown in Ref. [7], the two representations (8.8) and (8.10) have a common domain of validity: values of $\mathcal{M}$ and L for which $V\Sigma\mathcal{M}$ is large, while ML is small. The comparison of the two representations allows one to pin down the normalization constant occurring in Eq. (8.10):

$$A = \kappa \left[\frac{F^2 V^{1/2} e^{\beta_0}}{2\pi} \right]^{(N_f^2-1)/2} \; . \tag{8.11}$$

The constant β_0 is one of the shape coefficients associated with the kinematics of free particles on a torus [9] (for a torus of equal sides, $\beta_0 = 1,70, \ldots$). The factor κ in front arises from the Haar measure. Parametrizing the elements $U \in SU(N_f)$ by coordinates α^i, the group metric is defined by

$$g_{ik}(\alpha) d\alpha^i d\alpha^k = \tfrac{1}{2}\mathrm{tr}(dU\, dU^\dagger) \; . \tag{8.12}$$

The Haar measure is proportional to the square root of $g = \det g_{ik}$:

$$d\mu(U) = \frac{1}{\kappa}\sqrt{g} \prod_{i=1}^{N_f^2-1} d\alpha^i \; . \tag{8.13}$$

The factor κ is the proportionality constant occurring in this relation. If the volume $\int_{SU(N_f)} d\mu(U)$ is normalized to 1, κ is given by [22]

$$\kappa = \frac{(2\pi^{N_f+2})^{(N_f-1)/2}}{\prod_{k=1}^{N_f-1} k!} \sqrt{N_f} \; . \tag{8.14}$$

In the context of the present paper, the representation (8.10) plays a role analogous to Eq. (3.1), generalizing this formula to the case of two or more flavors. To our knowledge the group integral cannot be expressed in terms of known functions if $N_f \geq 3$. In the case of two flavors, however, the integration can be done and leads to a Bessel function of imaginary argument,

$$Z = \frac{2}{V\Sigma m_0} I_1(V\Sigma m_0) \quad (N_f = 2) \; , \tag{8.15}$$

where m_0 refers to the maximum in the integrand:

$$\begin{aligned} m_0 &= \max \mathrm{Re}\{e^{i\theta/2}\mathrm{tr}(\mathcal{M}U_0^\dagger)\} \quad [U_0 \in SU(2)] \\ &= \{\mathrm{tr}\mathcal{M}^\dagger\mathcal{M} + e^{i\theta}\det\mathcal{M} + e^{-i\theta}\det\mathcal{M}^\dagger\}^{1/2} \; . \end{aligned} \tag{8.16}$$

In a basis where $\mathcal{M}$ is diagonal, the quark condensate becomes

$$\begin{aligned} \langle \bar{u}_L u_R \rangle &= -\frac{1}{2}\Sigma \frac{m_u + e^{-i\theta} m_d^*}{|m_u + e^{-i\theta} m_d^*|} \frac{I_2(V\Sigma m_0)}{I_1(V\Sigma m_0)} \; , \\ \langle \bar{u}_R u_L \rangle &= \langle \bar{u}_L u_R \rangle^* \; , \end{aligned} \tag{8.17}$$

and similarly for $\langle \bar{d}_L d_R \rangle$, $\langle \bar{d}_R d_L \rangle$. The result shows that, in the chiral limit at fixed volume, the condensate disappears as it should—symmetries do not break down spontaneously if the volume is finite. In the symmetry-restoration region, $V\Sigma m_0 \ll 1$, the condensate is proportional to the quark mass and volume:

$$\langle \bar{u}_L u_R \rangle \simeq -\tfrac{1}{8}\Sigma^2 V(m_u + e^{-i\theta} m_d^*) \; . \tag{8.18}$$

The term $\propto m_u$ on the RMS stems from the nonzero mode contribution in the sector with $\nu = 0$, while the term $\propto m_d^*$ is due to zero modes in the sector with $\nu = -1$. If $m_u \sim m_d$, they are of the same order. On the other hand, for $V\Sigma m_0 \gg 1$, the condensate becomes volume independent. The magnitude of $\langle \bar{u}_L u_R \rangle$ tends to $\frac{1}{2}\Sigma$, while the phase is given by the phase of $m_u + e^{-i\theta} m_d^*$, irrespective of the volume [23,15]. Note the qualitative difference to the case $N_f = 1$, where the phase of the condensate is independent of the phase of the quark mass. For $N_f = 2$, we are dealing with a spontaneously broken symmetry and the quark condensate is of the same physical nature as the spontaneous magnetization of a ferromagnet, the mass term in the Lagrangian playing a role analogous to an external magnetic field. The minimum of the free energy is achieved if the magnetization is parallel to the external magnetic field. If the latter changes sign, so does the former. The phase of $\det\langle \bar{q}_R q_L \rangle$, however, is independent of $\arg\det\mathcal{M}$: The product $\langle \bar{u}_L u_R \rangle \langle \bar{d}_L d_R \rangle e^{i\theta}$ is positive—this is what remains of the mass independence of the condensate when going from $N_f = 1$ to 2.

If the two quark masses are real and if $\theta = 0$, the effective Lagrangian in Eq. (8.4) only involves the sum $m_u + m_d$. The isospin breaking generated by $m_u - m_d$ does not manifest itself at the leading order of the chiral expansion. The isospin symmetry remains unbroken in the leading order also in the more general case where the masses are complex and θ does not vanish. Setting $U(x) = U_0 U_1(x)$, where U_0 is the minimum of the effective action, the mass term reduces to $\Sigma m_0 \mathrm{Re}\,\mathrm{tr} U_1$. The effective Lagrangian is therefore invariant under isospin rotations, $U_1 \to V U_1 V^\dagger$. In particular, the masses of the three Goldstone bosons coincide:

$$M_\pi^2=\frac{\Sigma m_0}{F^2}=\frac{\Sigma}{F^2}|m_u+e^{-i\theta}m_d^*| \ . \qquad (8.19)$$

In a basis where m_u and m_d are positive, this reduces to Eq. (8.1) if θ vanishes, while for $\theta=\pi$ the pion mass is proportional to $|m_u-m_d|^{1/2}$. In particular, for $m_u=m_d$, $\theta=\pi$ (or $m_u=-m_d$, $\theta=0$), the effective Lagrangian in Eq. (8.4) describes three massless pions. In this case it is not legitimate to neglect the higher-order terms in the effective Lagrangian. The expansion of M_π^2 in powers of $\mathcal{M}$ then only starts at order $\mathcal{M}^2$. In the context of the present paper, these peculiarities are not of importance, because they only arise for particular values of the vacuum angle and do therefore not show up in the Fourier coefficients Z_ν.

Although the group integral in Eq. (8.10) cannot be done explicitly[5] for $N_f\geq 3$, approximate representations valid if $V\Sigma\mathcal{M}$ is either large or small compared to 1 are readily obtained. For $V\Sigma\mathcal{M}\gg 1$, the partition function grows exponentially with the volume, $Z\sim\exp(V\Sigma m_0)$. The quantity m_0 is given by the maximum of the exponent in Eq. (8.10). Denoting the group element at which this maximum occurs by U, the condensate becomes

$$\langle \bar q_L^{\,j} q_R^{\,i}\rangle \simeq -\tfrac{1}{2}\Sigma U^{ij}e^{-i\theta/N_f} \ . \qquad (8.20)$$

In the opposite extreme $V\Sigma\mathcal{M}\ll 1$, the exponential in Eq. (8.10) can be expanded in powers of the volume. Using the orthogonality relation for the fundamental representation of SU(N_f), this gives

$$Z=1+\frac{1}{4N_f}V^2\Sigma^2\mathrm{tr}(\mathcal{M}^\dagger\mathcal{M})+\cdots \quad (N_f\geq 3) \ . \qquad (8.21)$$

In this expansion the vacuum angle only shows up at order V^{N_f}, through a term proportional to $\mathrm{Re}\{e^{i\theta}\det\mathcal{M}\}$. The representation (8.21) shows that, in the symmetry-restoration region, the quark condensate is proportional to the mass matrix:

$$\langle \bar q_L^{\,j} q_R^{\,i}\rangle = -\frac{V\Sigma^2}{4N_f}\mathcal{M}^{ij}+\cdots \quad (N_f\geq 3) \ . \qquad (8.22)$$

It is due *exclusively* to nonzero modes in the sector with $\nu=0$.

IX. SPECTRUM OF THE DIRAC OPERATOR FOR $N_f\geq 2$

We now turn to the consequences of the representation (8.10) for the distribution of the winding number and fermionic eigenvalues. The contribution to the partition function generated by field configurations of winding number ν is given by the Fourier coefficient Z_ν. In the representation (8.10), the vacuum angle θ and collective meson-field variable U_0 only enter through the combination $U=U_0\exp(-i\theta/N_f)$. The Fourier transform with respect to θ therefore merely extends the group integral from SU(N_f) to U(N_f):

$$Z_\nu=A\int_{\mathrm{U}(N_f)}d\mu(U)(\det U)^\nu\exp\{V\Sigma\,\mathrm{Re}\,\mathrm{tr}(\mathcal{M}U^\dagger)\} \ . \qquad (9.1)$$

Here and in the following, the volume of the group $\int_{\mathrm{U}(N_f)}d\mu(U)$ is normalized to 1. For $N_f=1$, this integration can explicitly be done, yielding the modified Bessel function encountered in the first part of the present paper. For $N_f\geq 2$, the function Z_ν depends on a matrix variable

$$X=V\Sigma\mathcal{M} \qquad (9.2)$$

rather than on a single argument, but the qualitative properties are very similar to those of the Bessel function. In fact, the essence of the following discussion is the statement that the properties of Z_ν and the fermionic spectrum are nearly independent of the number of flavors. This is in marked contrast with the result of the preceding section where we found that the behavior of the full partition function and quark condensate strongly depends on the number of flavors, because, for $N_f\geq 2$, the theory acquires a symmetry which undergoes spontaneous breakdown. Remarkably, the restriction to field configurations of a given winding number removes the qualitative difference between the cases $N_f=1$ and $N_f\geq 2$.

Let us now establish the main properties of the group integral in Eq. (9.1). Exploiting the invariance of the Haar measure, one readily shows that Z_ν obeys the identity

$$Z_\nu(V_1\mathcal{M}V_2)=(\det V_1V_2)^\nu Z_\nu(\mathcal{M}) \ , \qquad (9.3)$$

valid for arbitrary $V_1,V_2\in U(N_f)$. Using a transformation of this type, the mass matrix can be brought to diagonal form with positive eigenvalues. It therefore suffices to analyze the properties of the integral in this case.

Let us first consider the region $V\Sigma\mathcal{M}\gg 1$, where the integrand in Eq. (9.1) is sharply peaked around the maximum occurring at $U=1$. Unless ν is taken as large, Z_ν therefore becomes independent of ν. As the winding number grows, the factor $(\det U)^\nu$ starts playing a role, reducing the value of the integral. In the region we are considering here, $V\Sigma\mathcal{M}\gg 1$, the winding-number distribution is Gaussian:

$$Z_n=Z(\theta=0)\frac{1}{(2\pi\langle\nu^2\rangle)^{1/2}}\exp\left[-\frac{1}{2}\frac{\nu^2}{\langle\nu^2\rangle}\right] \ . \qquad (9.4)$$

Since the mean-square winding number $\langle\nu^2\rangle$ is large, the sum over ν coincides with the corresponding integral. We have normalized the Gaussian distribution accordingly, such that the factor in front of this distribution coincides with $\sum_\nu Z_\nu$, i.e., with the partition function at zero vacuum angle.

The above derivation of Eq. (9.4) relies on the representation (9.1) for Z_ν. As discussed in the last section, this representation only holds if the box is small compared to the Compton wavelength of the Goldstone bosons. For $M_iL\gtrsim 1$, the behavior of the partition function is described by Eq. (8.8) rather than (8.10). To calculate the corresponding Fourier coefficients Z_ν, we observe that the vacuum-energy density occurring in formula (8.8) has

[5]See, however, the next section where, for quarks of equal mass, the Fourier coefficient Z_ν is given in closed form.

its minimum at $\theta=0$ (recall that we are taking $\mathcal{M}$ real and positive). The meson masses M_i contained in the function g_0 also depend on θ, but this term only represents a correction to the leading term ϵ_0. In the region $V\Sigma\mathcal{M}\gg 1$, the minimum in the vacuum energy generates a sharp peak in the function $Z=Z(\theta)$, described by a Gaussian in θ. The corresponding Fourier transform is therefore also a Gaussian: The winding-number distribution is given by Eq. (9.4), irrespective of the size of M_iL. Moreover, we can read off the mean-square winding number by calculating the ratio $Z''(\theta)/Z(\theta)$ at $\theta=0$. Up to higher-order corrections, this gives

$$\langle \nu^2\rangle = V\frac{\partial^2\epsilon_0}{\partial\theta^2}(\mathcal{M}e^{i\theta})|_{\theta=0} \; . \tag{9.5}$$

We have discussed the θ dependence of the vacuum energy in detail in Sec. VIII, for the case of two flavors, where

$$\epsilon_0 = -\Sigma(m_u^2+m_d^2+2m_u m_d\cos\theta)^{1/2} \; . \tag{9.6}$$

Evaluating the derivative, we obtain

$$\langle \nu^2\rangle = V\Sigma\frac{m_u m_d}{m_u+m_d} \; . \tag{9.7}$$

In fact, the behavior of the topological susceptibility $\langle\nu^2\rangle/V$ at infinite volume was analyzed in Ref. [23], where it is also shown that the result (9.7) can be generalized to an arbitrary number of flavors. The topological susceptibility is determined by the sum of the inverse quark masses:

$$\langle \nu^2\rangle = V\Sigma\left\{\frac{1}{m_u}+\frac{1}{m_d}+\frac{1}{m_s}+\cdots\right\}^{-1} . \tag{9.8}$$

Let us now turn to the opposite limit $V\Sigma\mathcal{M}\ll 1$, where symmetry restoration occurs. In this region the integrand in the group integral (9.2) does not vary rapidly and can be expanded in powers of $\mathcal{M}$ and $\mathcal{M}^\dagger$. The identity (9.3) implies that for $\nu\geq 0$ the expansion starts with a term proportional to $(\det\mathcal{M})^\nu$:

$$Z_\nu = \mathcal{N}_\nu(\det X)^\nu\{1+a_\nu \mathrm{tr}X^\dagger X + b_\nu(\mathrm{tr}X^\dagger X)^2 + c_\nu \mathrm{tr}X^\dagger X X^\dagger X + O(X^6)\} \; . \tag{9.9}$$

Evidently, the factor $(\det X)^\nu$ represents the contribution from the νN_f fermionic zero modes [for $\nu<0$, this factor is to be replaced by $(\det X^\dagger)^{-\nu}$].

We proceed with an evaluation of the coefficients a_ν, b_ν, c_ν—the overall normalization constant $\mathcal{N}_\nu$ will be determined later on. Decompose the matrix X as

$$X=\sum_{a=1}^{N_f^2} X^a t_a \; , \tag{9.10}$$

where $t_1,\ldots,t_{N_f^2}$ are the generators of $U(N_f)$ and constitute a complete set of Hermitian $N_f\times N_f$ matrices, normalized by

$$\mathrm{tr}(t_a t_b)=\tfrac{1}{2}\delta_{ab} \; , \qquad \sum_a t_a t_a = \tfrac{1}{2}N_f \; . \tag{9.11}$$

It is convenient to treat the variable X^a and their complex conjugates as independent and to denote the corresponding partial derivatives by ∂_a and $\bar\partial_a$, respectively. Using the identity

$$\sum_a \mathrm{tr}(t_a A)\mathrm{tr}(t_a B)=\tfrac{1}{2}\mathrm{tr}(AB) \; , \tag{9.12}$$

one readily checks that the integral (9.1) obeys the second-order differential equation

$$\sum_a \partial_a\bar\partial_a Z_\nu = \tfrac{1}{8}N_f Z_\nu \; . \tag{9.13}$$

This relation generalizes the familiar differential equation obeyed by the Bessel functions which corresponds to the case $N_f=1$. Inserting the expansion (9.9), one obtains

$$a_\nu = \frac{1}{4(N_f+|\nu|)} \; , \tag{9.14}$$

as well as a relation involving a particular linear combination of the coefficients b_ν and c_ν:

$$(N_f^2+|\nu|N_f+1)b_\nu + (2N_f+|\nu|)c_\nu = \frac{N_f}{32(N_f+|\nu|)} \; . \tag{9.15}$$

To separately determine b_ν and c_ν, one may use the fourth-order differential equation

$$\sum_{abcd}\mathrm{tr}(t_a t_b t_c t_d)\partial_a\bar\partial_b\partial_c\bar\partial_d Z_\nu = \tfrac{1}{16}N_f Z_\nu \; , \tag{9.16}$$

which also follows from the representation (9.1) upon repeated use of Eq. (9.12). Inserting the expansion (9.9) and using the identity

$$\sum_a t_a A t_a = \tfrac{1}{2}\mathrm{tr}A \; , \tag{9.17}$$

which holds for any $N_f\times N_f$ matrix A, this gives

$$(2N_f+|\nu|)b_\nu + (N_f^2+|\nu|N_f+1)c_\nu = \frac{1}{32(N_f+|\nu|)} \; . \tag{9.18}$$

Solving the linear system of equations (9.15) and (9.18), we finally obtain

$$b_\nu = \frac{1}{32(k^2-1)} \; , \qquad c_\nu = -\frac{1}{32k(k^2-1)} \; , \tag{9.19}$$

with $k\equiv|\nu|+N_f$. As a check, one may verify that for $N_f=1$ the representation (9.9) agrees with the first few terms of the Taylor series for the Bessel function.

The implications of these results for the spectrum of the Dirac operator are the following. The exponential growth of Z_ν at large values of $V\Sigma\mathcal{M}$ implies that, in the range $1/V\Sigma \ll |\lambda_n| \ll \Lambda_{QCD}$, the number of levels contained in the interval $\Delta\lambda$ is the same as for $N_f=1$:

$$\Delta n = \frac{1}{\pi}V\Sigma\,\Delta\lambda \; . \tag{9.20}$$

Note that to compare theories with a different number of light flavors, we need to relate the corresponding mass scales. In the present context, it is convenient to choose the scale such that the vacuum expectation value of one of the quark flavors, $\langle 0|\bar{u}u|0\rangle = -\Sigma$, is the same. In the range of eigenvalues specified above, the level density then becomes independent not only of the winding number, but also of the number of flavors.

For $N_f = 1$, we found that at the lower end of the spectrum, $\lambda_n = O(1/V\Sigma)$, the eigenvalue distribution does depend on the winding number. It also depends on the number of flavors. To see this one compares the expansion (9.9) with the path-integral representation (2.8) and (2.9). The comparison leads to a generalization of the sum rules established in Sec. VI. Expressed in terms of the dimensionless eigenvalues $l_n \equiv V\Sigma\lambda_n$, the sum rules take the form

$$\left\langle\!\!\left\langle \sum_n{}' \frac{1}{l_n^2} \right\rangle\!\!\right\rangle_\nu = \frac{1}{4k} , \tag{9.21a}$$

$$\left\langle\!\!\left\langle \left[\sum_n{}' \frac{1}{l_n^2}\right]^2 \right\rangle\!\!\right\rangle_\nu = \frac{1}{16(k^2-1)} , \tag{9.21b}$$

$$\left\langle\!\!\left\langle \sum_n{}' \frac{1}{l_n^4} \right\rangle\!\!\right\rangle_\nu = \frac{1}{16k(k^2-1)} , \tag{9.21c}$$

with $k \equiv |\nu| + N_f$. The mean value $\langle\langle \cdots \rangle\rangle_\nu$ is defined in Eq. (6.2), except that the fermionic factor occurring there is now given by $(\prod_n' \lambda_n^2)^{N_f}$,

The first relation shows that, in the presence of several flavors, the factor $|\nu|+1$ occurring in Eq. (6.10a) is replaced by $|\nu| + N_f$: The sum $\sum_n' 1/l_n^2$ decreases in inverse proportion to the sum of the winding number and number of flavors, indicating that the eigenvalue spectrum is pushed up if either one of these two quantum numbers is increased. Note that this statement refers to the dimensionless quantities l_n, which differ from the eigenvalues λ_n by the factor $V\Sigma$. Both the Pauli principle and the fact that the fermions tend to shield the attraction generated by the gluons indicate that the condensate Σ decreases if additional flavors are introduced at fixed coupling. This effect amplifies the shift toward larger eigenvalues caused by an increase in N_f. Qualitatively, the sensitivity to N_f is to be expected: The path integral for Z_ν contains a factor from the action of the gluon field and one from the determinant of the Dirac operator, the zero modes being factored out. The determinant grows if the spectrum is pushed toward larger eigenvalues. If the number of flavors is increased, the weight of the fermionic factor becomes more important such that gluon-field configurations with larger eigenvalues are favored.

For $N_f = 1$, only the difference between the two inverse moments in Eqs. (9.21b) and (9.21c) obeys a sum rule [see Eq. (6.10b)]. If there are several flavors, we thus obtain an additional piece of information which concerns the dispersion of the quantity $s \equiv \sum_n' 1/l_n^2$:

$$\frac{\langle\langle (s-\bar{s})^2 \rangle\rangle_\nu}{\bar{s}^2} = \frac{1}{k^2-1} , \tag{9.22}$$

where $\bar{s} \equiv \langle\langle s \rangle\rangle_\nu$. The result shows that the inverse moment s fluctuates around the mean value $\bar{s} = 1/4k$. The relative dispersion is independent of the volume, but decreases if the number of flavors or the winding number are increased.

Up to now, we have analyzed the partition function Z_ν for a general mass matrix $\mathcal{M}$. This allowed us to distinguish between the two fourth-order terms in the expansion (9.9) and to calculate separately the coefficients b_ν and c_ν which determine the two independent moments in Eq. (9.21b) and (9.21c). It is amusing, however, to note that, if the quark masses are set equal, the group integral Z_ν can be worked out explicitly for any number of flavors. We briefly sketch the calculation, dropping the overall normalization factor A to simplify the notation.

Consider first the case of two flavors. If the matrix X is proportional to the unit matrix, $X = x1$, the integral in Eq. (9.1) is readily evaluated, with the result

$$Z_\nu = I_\nu^2(x) - I_{\nu+1}(x) I_{\nu-1}(x) . \tag{9.23}$$

This expression represents a determinant. It suggests that, for an arbitrary number of flavors of equal mass, the partition function is given by[6]

$$Z_\nu = \begin{vmatrix} I_\nu(x) & I_{\nu+1}(x) & \cdots & I_{\nu+N_f-1}(x) \\ I_{\nu-1}(x) & I_\nu(x) & \cdots & I_{\nu+N_f-2}(x) \\ \vdots & \vdots & \vdots & \vdots \\ I_{\nu-N_f+1}(x) & \cdots & \cdots & I_\nu(x) \end{vmatrix} , \tag{9.24}$$

where x is related to the quark mass by $x = V\Sigma m$. To verify that this formula is correct, we first note that, for equal quark masses, the integrand in Eq. (9.1) is invariant under $U \to VUV^\dagger$ with $V \in U(N_f)$. The matrix U can therefore be diagonalized. The eigenvalues are of the form $e^{i\phi_1}, \ldots, e^{i\phi_{N_f}}$. Using Weyl's formula [24], the group integral can be reduced to an integral over the angles $0 \le \phi_1, \ldots, \phi_N \le 2\pi$:

$$Z_\nu = \frac{1}{N_f!} \int \prod_{i=1}^{N_f} \frac{d\phi_i}{2\pi} |P|^2 \exp\left[\sum_{k=1}^{N_f} \{x \cos\phi_k + i\nu\phi_k\}\right] , \tag{9.25}$$

where P is the function

$$P = \prod_{k<l} (e^{i\phi_k} - e^{i\phi_l}) . \tag{9.26}$$

If P is multiplied out, we obtain a set of terms for which the integral factorizes into N_f independent one-dimensional integrals, each of which representing a Bessel function. To check that the combinatorics of these terms indeed gives rise to the determinant (9.24), we note that P is a linear combination of factors of the form $\exp(i\{n_1\phi_1 + \cdots + n_{N_f}\phi_{N_f}\})$. Since the result must be odd under the interchange of any two angles ϕ_i, ϕ_k, only those terms survive for which the integers $(n_1, \ldots, n_{N_f})$ are mutually different. Hence the set of integers must

[6]We are indebted to H. B. Nielsen for this hint.

represent a permutation of the numbers $(0,1,2,\ldots,N_f-1)$; i.e., P consists of $N_f!$ terms which only differ by a permutation of $\phi_1,..,\phi_{N_f}$ and by a sign if the permutation is odd. Now the integrand in Eq. (9.25) involves the factor PP^*. In view of the permutation symmetry of the remainder, we can replace the quantity P^* by a single term, say, $(n_1,\ldots,n_{N_f})=(0,1,\ldots,N_f-1)$, dropping the factor $1/N_f!$ in front of the integral. One then easily checks that the $N_f!$ permutations occurring in P precisely give rise to the determinant (9.24).

At small x the partition function (9.24) is proportional to $x^{|\nu|N_f}$ as is dictated by the general representation (9.9) reflecting the presence of $|\nu|N_f$ fermionic zero modes. The proportionality coefficient can also be worked out (see Appendix B):

$$Z_\nu=\left[\frac{x}{2}\right]^{|\nu|N_f}\sum_{k=0}^{N_f-1}\frac{k!}{(k+|\nu|)!}[1+O(x^2)] \quad (x\ll 1)\ . \tag{9.27}$$

Reinserting the overall normalization factor A given in Eq. (8.11), this finally allows us to pin down the constant $\mathcal{N}_\nu$ occurring in the small X expansion (9.9):

$$\mathcal{N}_\nu=\sqrt{N_f}\,\frac{2^{-N_f(|\nu|+N_f/2-1/2)}\pi^{(N_f-1)/2}}{\prod_{k=0}^{N_f-1}(k+|\nu|)!}\times(F^2V^{1/2}e^{\beta_0})^{(N_f^2-1)/2}\ . \tag{9.28}$$

X. ADJOINT FERMIONS

The preceding analysis concerns the QCD sector of the standard model, where the fermions transform according to the fundamental representation of the gauge group. It is instructive to see what happens if the fermions of the theory instead sit in the adjoint representation. Supersymmetric Yang-Mills (SYM) theory is an example of this category, containing one massless Majorana fermion in the adjoint representation of the gauge group.

The cardinal difference to the case discussed in the preceding sections is that, if the theory exclusively involves particles in the adjoint representation, the winding number defined in Eqs. (2.2) and (2.6) need not be an integer. To see why this is so, we recall that the quantization of the winding number in QCD is a consequence of the requirement that the gauge field must be able to support quarks. For fermions in the adjoint representation $D(g)$ of $SU(N_c)$, the periodicity condition (2.4) is replaced by

$$\psi(x+a)=(-1)^{|a|}D[\Omega_a(x)]\psi(x)\ . \tag{10.1}$$

Accordingly, the consistency condition (2.5) for the transition function now becomes

$$D[\Omega_{a+b}(x)]=D[\Omega_b(x+a)]D[\Omega_a(x)]\ . \tag{10.2}$$

The point is that this represents a weaker condition. It holds if and only if $\Omega_a(x)$ obeys [25]

$$\Omega_{a+b}(x)=Z(a,b)\Omega_b(x+a)\Omega_a(x)\ , \tag{10.3}$$

where $Z(a,b)$ is an element of the center of $SU(N_c)$. [The adjoint representation of the center is trivial, $D(Z)=1$.] It remains true that the winding numbers of the gauge field can be expressed in terms of the transition function, but if the composition rule (10.3) involves nontrivial Z factors, ν is in general a fractional number.

The admissible winding numbers can be determined as follows. Since in the adjoint representation the covariant derivative

$$(D_\mu)^{bc}=\delta^{bc}\partial_\mu+iG_\mu^a(T^a)^{bc}=\delta^{bc}\partial_\mu+f^{abc}G_\mu^a \tag{10.4}$$

is real, the eigenvalue equation $\not{D}u_n=\lambda_n u_n$ is invariant under charge conjugation: If $u_n(x)$ is an eigenfunction, then $C^{-1}u_n^*(x)$ is also an eigenfunction and the eigenvalue λ_n is the same (C is the charge-conjugation matrix defined by $\gamma_\mu^*=-\gamma_\mu^T=C\gamma_\mu C^{-1}$). In view of $C^*C=-1$, the two solutions are linearly independent. Hence the eigenfunctions occur in pairs:

$$\{\lambda_n:u_n(x),C^{-1}u_n^*(x)\}\ . \tag{10.5}$$

In particular, the number of zero modes is even. Denote the number of left- and right-handed zero modes by n_L and n_R, respectively. Since charge conjugation preserves chirality, both n_L and n_R are even integers. Now, according to the index theorem, the difference n_L-n_R is given by the integral $(16\pi^2)^{-1}\int d^4x\,\mathrm{tr}(G_{\mu\nu}\tilde{G}^*_{\mu\nu})$, where the matrix $G_{\mu\nu}$ is the adjoint representation of the field strength. This integral must therefore be an even integer. The trace over the generators of the adjoint representation differs from the analogous trace for the fundamental representation by the factor $2N_c$. The above integral thus coincides with $2\nu N_c$, where ν is the winding number defined in Eqs. (2.2) and (2.6). Hence a necessary condition for a gauge-field configuration to support adjoint fermions is that the winding number be of the form

$$\nu=\frac{\bar{\nu}}{N_c}\ , \tag{10.6}$$

where $\bar{\nu}\equiv\frac{1}{2}(n_L-n_R)$ is an integer. To show that this condition on the winding number is also sufficient, one may, e.g., consider Abelian configurations of constant field strength, referred to as torons [26]. In that case the transition functions can be worked out explicitly and one readily verifies that the condition (10.3) is indeed obeyed if and only if the winding number is of the fractional form (10.6). Note that self-dual or anti-self-dual configurations of this type only generate a subset of the admissible winding numbers.

Let us now see what becomes of our analysis of the partition function. Since the adjoint representation is real, it is convenient to describe the fermions in terms of Majorana fields. The Lagrangian of our model then reads

$$\mathcal{L}=\frac{1}{4g^2}G_{\mu\nu}^aG_{\mu\nu}^a-i\theta\omega-i\sum_i\bar{\lambda}^i\not{D}\lambda^i+\sum_{ik}\{m_{ik}\bar{\lambda}_R^i\lambda_L^k+\mathrm{H.c.}\}\ , \tag{10.7}$$

where D_μ and ω are defined in Eqs. (10.4) and (2.2), respectively. We denote the number of Majorana flavors by $\bar{N}_f$.

Before proceeding further, we should discuss a known problem which appears when one tries to define the functional integral in a theory involving Majorana (or Weyl) fermions in Euclidean space. The matter is related to the fact that the Euclidean rotation group SO(4) consists of two independent SU(2) factors—their generators are not Hermitian conjugates of one another as in the case of the Lorentz group. This means that a left Weyl spinor and its complex conjugate transform with the same representation of SO(4). As a result, it is impossible to construct an SO(4)-invariant Lagrangian involving only left Weyl fields—necessarily, fields with the opposite chirality appear and the number of degrees of freedom is effectively doubled [27]. By the same token, Majorana fermions which are normally composed of a Weyl spinor and its complex conjugate do not live in Euclidean space: In view of $C^*C=-1$, the Majorana condition

$$\lambda^i(x)=C\lambda^i(x)^* \tag{10.8}$$

immediately leads to a contradiction.

Although Majorana fields as such cannot be defined in Euclidean space, it is possible to analyze the theory characterized by the Lagrangian (10.7) by means of analytic continuation from Minkowski space [28]. The point is that, in Minkowski space, where it is well defined, the path integral over the Majorana fields can be done. It is given by the square root of the corresponding Dirac determinant. The determinant as such does have an unambiguous continuation to Euclidean space. Taking square roots of determinants in general meets with difficulties [29], but for Majorana fermions, there is no problem. As noted above, the eigenfunctions of $\not{D}$ occur in pairs related by charge conjugation. Furthermore, multiplication with γ_5 generates a pair of eigenfunctions belonging to the eigenvalue $-\lambda_n$. The square root can therefore explicitly be taken, and the formula analogous to Eq. (2.9) becomes

$$[\det(-i\not{D}+\tilde{\mathcal{M}})]^{1/2}=(\det{}_f\mathcal{M})^{\bar{\nu}}\prod_n{}''\det{}_f(\lambda_n^2+\mathcal{M}\mathcal{M}^\dagger)\ . \tag{10.9}$$

Again, the product runs over all positive eigenvalues, but now each pair of charge-conjugate eigenfunctions only counts once. The degeneracy (10.5), which allowed us to take the square root, is special for the adjoint representation. If the fermions belong to the fundamental representation, then, generally, T^{a*} is not related to T^a and the symmetry is absent. The exception is SU(2), where τ_a^* and τ_a are related by the unitary transformation τ_2: For any eigenfunction u_n of the corresponding Dirac equation, the function $\tau_2C^{-1}u_n^*$ is also an eigenfunction with the same eigenvalue. Even in this case, the solutions u_n and $\tau_2C^{-1}u_n^*$ need not be linearly independent—the equation $u_n=\tau_2C^{-1}u_n^*$ can have solutions.

We are now able to define the partition function of the theory, $Z(\theta)$, as a sum over the admissible winding numbers $\nu=0,\pm1/N_c,\pm2/N_c,\dots$. Each sector comes with the fermionic factor specified in Eq. (10.9), multiplied by $e^{i\nu\theta}$. The prescription implies, in particular, that the mass matrix $\mathcal{M}$ and vacuum angle θ only enter in the combination $\mathcal{M}\exp(i\theta/\bar{N}_fN_c)$. Since fractional winding numbers occur, $Z(\theta)$ is periodic in θ with period $2\pi N_c$ (not 2π as in the case of QCD).

We emphasize that the extended periodicity interval is inherent to our specific framework, where the Euclidean path integral is extended over *all* gauge-field configurations on a torus. As discussed in Sec. II, one may consider alternative states, described by different partition functions. One may, e.g., restrict the path integral to gauge fields of integer winding number. The corresponding partition function is given by

$$Z^*(\theta)=\sum_{n=1}^{N_c}Z(\theta+2\pi n) \tag{10.10}$$

and is periodic with period 2π. As long as the fermions are massive, the states $Z(0)$ and $Z^*(0)$ are indistinguishable in the infinite-volume limit. Note, however, that for strictly massless fermions (SYM theory), the vacuum energy becomes independent of θ. In this case the restriction to integer winding numbers does modify the state of the system, even at infinite volume. For $\bar{N}_f=1$, e.g., $Z(\theta)$ is the partition function of a system with a single ground state, while $Z^*(\theta)$ corresponds to N_c degenerate vacua [30]. In our opinion, the choice (10.10) is less natural than ours. More importantly, models involving massless fermions admit a variety of inequivalent partition functions. Nothing forbids, e.g., considering the sum $\sum_{n=1}^{2N_c}Z(\theta+n\pi)$, which is periodic in θ with period π and implies the occurrence of $2N_c$ degenerate vacua.

We now return to the analysis of the partition function for massive fermions where this problem does not arise. The simplest case, analogous to QCD with one quark flavor, occurs for $\bar{N}_f=1$ (for vanishing mass, this is the supersymmetric Yang-Mills theory). The vector current then vanishes identically while the axial-vector current is anomalous. The corresponding Ward identity confirms the conclusion drawn from the number of fermionic zero modes occurring in the Euclidean path integral: The partition function depends on the Majorana mass m and vacuum angle θ only through the product $m\exp(i\theta/N_c)$. For $\bar{N}_f=1$, the mass spectrum of the theory contains a gap, such that, at large volume, the partition function grows exponentially:

$$Z=\exp\{-V\varepsilon_0(me^{i\theta/N_c})\}\ , \tag{10.11}$$

where $\varepsilon_0(me^{i\theta/N_c})$ is the energy density of the ground state. If the mass m is small, ε_0 is approximately linear in m and the properties of the partition function are therefore again determined by the Taylor coefficient σ, which represents the vacuum expectation value

$$\langle 0|\bar{\lambda}\lambda|0\rangle=-\sigma\ . \tag{10.12}$$

Obviously, the situation is the same as in QCD with one quark flavor, and the analysis of Secs. IV–VII goes through without essential modifications. If the fermion mass is small and if the scales of the two theories are chosen such that the vacuum expectation values are the same, $\langle 0|\bar{\lambda}\lambda|0\rangle=\langle 0|\bar{q}q|0\rangle$, the partition functions become identical, except that one needs to compare the vacuum angle of the Majorana theory with the angle θ/N_c in QCD and compare configurations of winding number ν in the Majorana case with those of winding number νN_c in QCD:

$$Z_\nu^{\text{adjoint}} = Z_{\nu N_c}^{\text{fundamental}} = I_{\nu N_c}(V\sigma m) \ . \tag{10.13}$$

In particular, at small values of $V\sigma m$, the condensate originates in field configurations with winding number $\nu = \pm 1/N_c$, which play a role analogous to the instanton-like configurations in QCD. Indeed, this was observed long ago [31] in the context of SYM theory. In an interesting recent paper. Zhitnitsky used more elaborate arguments to calculate the gluino condensate (10.12) explicitly [32]. His result is $\sigma = c\Lambda_{\text{SYM}}^3$, where Λ_{SYM} is the analogue of Λ_{QCD} and c is a known numerical coefficient.[7]

In the opposite limit $V\sigma m \gg 1$, the value of the condensate is the same. It does then not arise from zero modes (as is the case in the limit of small $V\sigma m$) but from small nonzero eigenvalues, and the winding number becomes an irrelevant quantum number.

If the theory contains several Majorana flavors $\lambda^1, \ldots, \lambda^{\bar{N}_f}$, the Lagrangian acquires additional chiral symmetries.[8] The currents

$$V_\mu^{ik} = \bar{\lambda}^i \gamma_\mu \lambda^k, \quad A_\mu^{ik} = \bar{\lambda}^i \gamma_\mu \gamma^5 \lambda^k \tag{10.14}$$

are conserved, except for the anomalous singlet current $\Sigma_i A_\mu^{ii}$. The kinematic identities $V_\mu^{ik} = -V_\mu^{ki}$, $A_\mu^{ik} = A_\mu^{ki}$ imply that we dealing with $\frac{1}{2}\bar{N}_f(\bar{N}_f - 1)$ and $\frac{1}{2}\bar{N}_f(\bar{N}_f + 1) - 1$ independent conserved vector and axial-vector currents, respectively. Accordingly, the symmetry group involves $\bar{N}_f^2 - 1$ parameters, and one readily checks that the commutation rules of the generators are those of $SU(\bar{N}_f)$. This is to be compared to the symmetry group $SU(N_f)_R \otimes SU(N_f)_L \otimes U(1)_V$, characteristic of N_f flavors in the fundamental representation.

The left-handed spinor $\lambda_L^i = \frac{1}{2}(1-\gamma^5)\lambda^i$ and the spinor $\bar{\lambda}_R^i$ transform according to the fundamental representation D_f of $SU(\bar{N}_f)$. The mass term $m_{ik}\bar{\lambda}_R^i \lambda_L^k$ therefore belongs to the direct product $D_f \otimes D_f$, more precisely, to the symmetric part thereof (the operators $\bar{\lambda}^i \lambda^k$ and $\bar{\lambda}^i \gamma^5 \lambda^k$ are symmetric under $i \leftrightarrow k$).

It is not evident that chiral symmetry undergoes spontaneous breakdown also if the fermions are in the adjoint representation. In the following we assume that it does, with a nonzero order-parameter matrix $\langle 0|\bar{\lambda}^i \lambda^k|0\rangle$, which breaks the $SU(\bar{N}_f)$ symmetry of the Lagrangian down to the subgroup $O(\bar{N}_f)$ generated by the vector currents. Accordingly, the theory must contain $\frac{1}{2}\bar{N}_f(\bar{N}_f+1) - 1$ Goldstone bosons which live in the quotient $SU(\bar{N}_f)/O(\bar{N}_f)$. Generally, the matrix $\langle 0|\bar{\lambda}^i \lambda^k|0\rangle$ depends on $\mathcal{M}$ and θ (cf. the discussion in Sec. VIII). Taking $\mathcal{M}$ diagonal with real positive entries and setting $\theta = 0$, the condensate at infinite volume takes the form

$$\langle 0|\bar{\lambda}^i \lambda^k|0\rangle = -\sigma\delta^{ik} \ , \tag{10.15}$$

up to corrections of order $\mathcal{M}$.

Let us now consider the case $\bar{N}_f = 2$, which amounts to a full Dirac spinor ψ in the adjoint representation of the gauge group. For simplicity, we restrict ourselves to a Dirac mass term of the form $m\bar{\psi}_R\psi_L + \text{H.c.}$ In the notation introduced above, this term corresponds to a mass matrix of the form $m_{ik} = m\delta_{ik}$ and breaks the SU(2) symmetry of the massless Lagrangian in the same manner as the condensate, with O(2) as residual symmetry. The quotient SU(2)/O(2) is a two-dimensional sphere. The Goldstone bosons can therefore be described by a vector $\mathbf{U}(x) = (U_1, U_2, U_3)$ of unit length, $\mathbf{U}^2 = 1$. The leading term in the effective Lagrangian is given by[9]

$$\mathcal{L}_{\text{eff}} = \tfrac{1}{2}f^2 \partial_\mu \mathbf{U} \partial_\mu \mathbf{U} - 2\sigma\, \text{Re}(me^{i\bar{\theta}})U_3 \ , \tag{10.16}$$

where we have identified the O(2) subgroup generated by the vector current with the rotations around the third axis. The angle $\bar{\theta}$ stands for $\theta/2N_c$. The representation of the partition function analogous to Eq. (8.10) now becomes

$$Z = \int \frac{d\Omega}{4\pi} \exp\{2V\sigma\, \text{Re}(me^{i\bar{\theta}})U_3\} \ , \tag{10.17}$$

where $d\Omega$ is the area element on the two-sphere. The angular integral can be done and leads to

$$Z = \frac{\sinh(2V\sigma m \cos\bar{\theta})}{2V\sigma m \cos\theta} \ , \tag{10.18}$$

where we have taken m to be real to simplify the notation. The corresponding Fourier coefficients are

$$Z_\nu = \sum_{n=0}^{\infty} \frac{x^{2n+2|\bar{\nu}|}}{n!(n+2|\bar{\nu}|)!(2n+2|\bar{\nu}|+1)} \ , \tag{10.19}$$

with $\bar{\nu} = N_c\nu$ and $x = V\sigma m$. The properties of the partition function (10.19) are rather similar to the corresponding expression (9.23) for the fundamental fermions. At $x \ll 1$, $Z_\nu \propto x^{2|\bar{\nu}|}$, and the contribution of large winding numbers is suppressed. At $x \gg 1$, the asymptotics for $Z(\bar{\theta})$ is exponential,

$$Z(\bar{\theta}) \propto \exp(2V\sigma m \cos\bar{\theta}) \ , \tag{10.20}$$

the Fourier integral for Z_ν is squeezed to small $\bar{\theta} \sim 1/\sqrt{x}$, and the contribution of all winding numbers up to $\bar{\nu} \sim \sqrt{x}$ is equally important.

---

[7]Note that, historically, the gluino condensate in SYM theory was first found in [33] by considering the instanton contribution to the correlator $0|T\{\bar{\lambda}\lambda(x), \ldots, \bar{\lambda}\lambda(0)\}|0\rangle$ [with N_c factors of $\bar{\lambda}\lambda(x_i)$].

[8]The first coefficient in the perturbative expansion of the β function is given by $\beta_0 = \frac{1}{3}N_c(11 - 2\bar{N}_f)$. The theory is therefore asymptotically free, provided there are less than six Majorana flavors.

[9]The model we are considering here belongs to the general class of $O(N)$-symmetric theories which undergo spontaneous breakdown to $O(N-1)$ [9]. We are dealing with the special case $N = 3$.

Note that, in full analogy with QCD, the expression (10.18) only holds in the region $ML \ll 1$, where $M=(2\sigma m \cos\theta/f^2)^{1/2}$ is the Goldstone-boson mass. For $ML \gtrsim 1$, chiral perturbation theory yields a representation analogous to Eq. (8.8). The corresponding large-x asymptotics is again dominated by the vacuum energy and coincides with Eq. (10.20)—the Goldstone modes only affect the preexponential factor.

The implications for the spectrum of the Dirac operator are also rather similar to what we had in the fundamental case. At $x \gg 1$, the condensate (10.15) is related to the mean density of the Dirac eigenvalues, $\rho(\lambda)$, per unit volume:

$$\langle 0|\bar{\lambda}^i\lambda^k|0\rangle = -\frac{\pi}{2}\rho(0)\delta^{ik} . \qquad (10.21)$$

Note the difference with the formula of Banks and Casher [Eq. (1.1)]. The factor of 2 arises because the density $\rho(\lambda)$ simply counts the eigenvalues of the Dirac operator, while for Majorana fermions only pairs of charge-conjugate eigenfunctions are relevant—the proper fermionic weight in the path integral is not the determinant but its square root.

Comparing the results (10.13) and (10.19) with $\langle\langle \det^{1/2}(-i\not{D}+\tilde{\mathcal{M}})\rangle\rangle_\nu$ for fermions of equal, positive mass m and performing the expansion in powers of $x=V\sigma m \ll 1$, we obtain the sum rules

$$\left\langle\!\!\left\langle \sum{}'' \frac{1}{\lambda_n^2} \right\rangle\!\!\right\rangle_\nu = \frac{V^2\sigma^2}{4(|\bar{\nu}|+1)} \quad (\bar{N}_f=1) , \qquad (10.22a)$$

$$\left\langle\!\!\left\langle \sum{}'' \frac{1}{\lambda_n^2} \right\rangle\!\!\right\rangle_\nu = \frac{V^2\sigma^2}{2(2|\bar{\nu}|+3)} \quad (\bar{N}_f=2) , \qquad (10.22b)$$

where the sums only run over pairs of charge-conjugate levels (10.5) with positive eigenvalues.

Let us compare these results with what we found in the case of QCD. Relation (10.13) implies that, in the presence of one Majorana flavor, the small eigenvalues of the Dirac operator are distributed in the same manner as in QCD with one quark flavor, except for two important differences. (i) We need to compare the eigenvalue distribution for gauge-field configurations of winding number ν in QCD to the distribution in the sector $N_c\nu$ of the Majorana theory. This means that the Majorana levels are more sensitive to the properties of the gluon field: Increasing the winding number from $\nu=0$ to 1, the sum $\sum_n \lambda_n^{-2}$ shrinks by a factor 2 in QCD ($N_f=1$), but by the factor N_c+1 in the Majorana case ($\bar{N}_f=1$). (ii) Tuning the scales of the two theories in such a way that the condensates are the same, $\sigma=\Sigma$, the eigenvalue distributions coincide at the lower end, but each level occurring in QCD corresponds to a pair of degenerate eigenvalues of the Dirac operator in the adjoint representation.

We add a comment concerning the significance of the scale convention $\sigma=\Sigma$ used in the above comparison of the two theories. The effective low-energy theory does not say anything about the magnitude of the low-energy constants $\Sigma, F, \sigma, f, \ldots$. In the large-N_c limit, the quark condensate Σ grows in proportion to N_c if the mass of one of the bound states, say, the ρ meson, is kept fixed [34]. In the case of adjoint fermions, on the other hand, the condensate σ is proportional to N_c^2. This implies that, if the scales of the two theories are tuned in such a manner that the value of M_ρ is the same, the condensates are of a different order of magnitude. At the lower end of the QCD spectrum, the spacing of the Dirac eigenvalues is of order $1/V\Sigma \propto 1/N_c$. In the Majorana case, the levels are more densely packed, with a spacing of order $1/V\sigma \propto 1/N_c^2$. The same difference in the level densities also shows up at large eigenvalues: For fermions in the fundamental representation, the number of levels contained in the interval $\Delta\lambda$ is given by $N_c V\lambda^3\Delta\lambda/4\pi^2$, while for adjoint fermions the factor N_c is to be replaced by N_c^2-1. The tuning of scales required by $\sigma=\Sigma$ therefore amounts to a comparison of the two theories at physically rather different scales, the bound states of the Majorana theory then occurring at masses which are large compared to those of the corresponding QCD states. Note also that, at leading order in an expansion in powers of $1/N_c$, the quark condensate Σ is predicted to become independent of the number of flavors, while for fermions in the adjoint representation this is not the case. Quite generally, theories containing adjoint fermions are more sensitive to the number of flavors than QCD. The sum rules only control the distribution of the dimensionless "eigenvalues" l_n. Comparison of the sum rules (10.22) and (9.21) shows that, in these variables, the flavor dependence of the two theories is rather similar. The factors $V\Sigma$ and $V\sigma$, which set the scale for the corresponding distributions of the proper eigenvalues λ_n, however, are quite different and also exhibit a different dependence on the number of flavors—it does not make much sense comparing the numerical coefficients which occur in the various sum rules.

XI. LARGE N_c

As pointed out in Ref. [35], the axial-vector anomaly is suppressed in the large-N_c limit. In this limit the theory therefore becomes symmetric under the full chiral group $U(N_f)_L \times U(N_f)_R$. There are good reasons to assume that this symmetry is spontaneously broken to the subgroup $U(N_f)_V$ generated by the vector currents [36]. Hence an additional Goldstone boson must occur if N_c is sent to infinity. The corresponding extension of the effective Lagrangian which describes the low-energy structure of the theory is discussed in detail in the literature [23,37,38]. Let us see what becomes of our analysis in this limit.

If the number of colors is large, the leading contributions to the partition function arise from graphs which exclusively involve gluons. In $\ln Z$, these graphs generate a contribution of order N_c^2, while graphs containing quarks only start showing up at order N_c. The sum of the purely gluonic contributions represents the partition function of gluodynamics ($N_f=0$).

Actually, this statement is not quite correct. In gluodynamics, the sum over *all* gluon-field configurations includes fractional winding numbers and the corresponding partition function is periodic in θ with period $2\pi N_c$, not 2π. In the presence of quarks, the path integral, however, only extends over the subset of field configurations with integer winding number. Denoting the full partition function of gluodynamics by $Z_G = Z_G(\theta)$, we are thus concerned with the superposition

$$\hat{Z}_G(\theta) = \frac{1}{N_c} \sum_{k=0}^{N_c-1} Z_G(\theta + 2\pi k) , \tag{11.1}$$

which eliminates fractional winding numbers, but retains integer ones with the proper weight. The difference between Z_G and $\hat{Z}_G$ clarifies a point which in the standard treatment of the large-N_c limit remains obscure. The point is the following. Gluodynamics is expected to have a mass gap M_0, given by the mass of the lightest glueball. Accordingly, if the box is large enough, $M_0 L \gg 1$, the excitations freeze and Z_G reduces to the contribution from the ground state:

$$Z_G(\theta) = e^{-V\epsilon_G(\theta)} . \tag{11.2}$$

The large-N_c counting rules imply that the dependence of the vacuum-energy density $\epsilon_G(\theta)$ on θ and the number of colors is of the form

$$\epsilon_G(\theta) = N_c^2 f\left[\frac{\theta}{N_c}\right] . \tag{11.3}$$

This property is perfectly consistent with periodicity on the interval $2\pi N_c$, but is difficult to understand, if, *ab initio*, the field configurations of gluodynamics are restricted to a integer winding number.

Note that the large-N_c limit can be replaced by the heavy quark limit at finite N_c—in either case, the partition function is given by the quantity $\hat{Z}_G(\theta)$, i.e., by the projection of the gluonic partition function onto integer winding numbers.

The path-integral representation shows that $Z_G(\theta)$ is a Fourier series with positive coefficients. On the interval $-\pi \le x \le \pi$, the function $f(x)$ therefore has an absolute minimum at $x=0$. For any given value of θ in the interval $-\pi \le \theta \le \pi$, the superposition (11.1) contains N_c terms, which populate the periodicity interval of the function $f(x)$ rather densely when N_c becomes large. In particular, irrespective of the value of θ, the superposition always contains terms for which $x = (\theta + 2\pi k)/N_c$ is close to the minimum. If the box is large compared to the scale of the theory, the quantity $\hat{Z}_G(\theta)$ is dominated by these terms.

For example, at $N_c = 2$, $\hat{Z}_G(\theta)$ involves two terms. At $|\theta| < \pi$, only the term with $k=0$ is relevant if the volume is large while, at $\pi < |\theta| < 2\pi$, the relevant term corresponds to $k=1$. The points $\theta = \pm\pi$ are special as both terms are there equally important; a change of regime occurs there. One can say that the system undergoes a phase transition at $\theta = \pm\pi$. The vacuum energy defined as $-V^{-1}\ln\hat{Z}_G(\theta)$ has a cusp singularity at these points.

The effects due to finite quark mass change the situation. In the case of a single quark flavor, these effects remove the singularity. Ths can be seen from the representation (4.2) valid for small quark mass. For several flavors, the singularity may not disappear in some cases. For example, for $N_f = 2$, the cusp persists at finite quark masses if they are all equal while, in the unequal mass case, the vacuum energy is a a smooth function of θ [15,23,43] [see also Eqs. (8.15) and (8.16)].

Invariance under parity implies that $f(x)$ is an even function of x. In the vicinity of the minimum, $f(x)$ is therefore of the form $f(0) + \frac{1}{2}x^2 f''(0) + O(x^4)$. The coefficent of the quadratic term, $\tau \equiv f''(0)$, represents the topological susceptibility of gluodynamics at $\theta = 0$:

$$\tau = \int dx \, \langle \omega(x)\omega(0) \rangle_G . \tag{11.4}$$

The corresponding mean-square winding number is proportional to the volume:

$$\langle \nu^2 \rangle_G = V\tau . \tag{11.5}$$

If the box is large, $V\tau \gg 1$, the superposition (11.1) is dominated by the term with $k=0$, the remainder only generating exponentially small corrections:

$$\hat{Z}_G = \frac{1}{N_c} e^{-V\epsilon_G(\theta)} . \tag{11.6}$$

Moreover, since we take N_c to be large, the vacuum-energy density is approximately quadratic in θ:

$$\epsilon_G(\theta) = N_c^2 f(0) + \tfrac{1}{2}\tau\theta^2 + O(\theta^4/N_c^2) . \tag{11.7}$$

In our context, where all sides of the torus are assumed to be large compared to the scale of the theory, the leading term of order N_c^2 is irrelevant and can be absorbed in the overall normalization constant, together with the factor $1/N_c$ in front of the exponential n Eq. (11.6). [At temperatures $T = 1/L_4$ of order $\Lambda_{\rm QCD}$, this term does contain interesting physics—in the large-N_c limit, it yields the leading contribution to the pressure.] We conclude that, although the purely gluonic graphs dominate the partition function of QCD in the large-N_c limit, they only generate a weak dependence on θ, determined by the topological susceptibility τ which is of order $(N_c)^0$. It is clear that $\hat{Z}_G$ does not provide an adequate representation of the QCD partition function to this accuracy—quark loops start contributing to $\ln Z$ already at order N_c.

The contributions to the partition function generated by the quark degrees of freedom can be analyzed in the same manner as at finite N_c, using an effective Lagrangian. For a derivation of the form of this Lagrangian, we refer the reader to the literature [23,37]. The main modification is that the spectrum of the theory now contains N_f^2 rather than $N_f^2 - 1$ light mesons. In the real world, the extra particle corresponds to the η', with $M_{\eta'} = 957$ MeV. In the following we will use this terminology independently of the number of flavors, referring to the meson which acquires Goldstone character only in the limit $N_c \to \infty$ as the η'. The Zweig rule which emerges from the large-N_c analysis implies that, in the chiral limit, the "pion-decay constant" of the η' is the same as for the ordinary Goldstone bosons, up to corrections of order $1/N_c$. It is therefore convenient to treat the N_f^2 particles on equal footing, in terms of a matrix field $U(x)$ which lives on the group $U(N_f)$. The η' is then described by the phase of the determinant:

$$\det U(x)=e^{-i\phi(x)} . \quad (11.8)$$

In this notation the effective Lagrangian which characterizes the low-energy structure of the theory in the large-N_c limit is of the form [23,37]

$$\mathcal{L}_{\rm eff}=\frac{F^2}{4}{\rm tr}(\partial_\mu U^\dagger \partial_\mu U) -\Sigma\, {\rm Re}\, {\rm tr}(\mathcal{M}U^\dagger)+\frac{\tau}{2}(\phi-\theta)^2 . \quad (11.9)$$

It involves three low-energy parameters: the pion-decay constant $F=O(N_c^{1/2})$, the parameter $\Sigma=O(N_c)$ which determines the quark condensate in the chiral limit at infinite volume, and the topological susceptibility of gluodynamics discussed above, τ of order 1.

The formula (11.9) represents an exact result in the following sense. Expand the full effective Lagrangian in powers of derivatives of the meson field and powers of the quark-mass matrix. Furthermore, expand the effective coupling constants which occur in this series in powers of $1/N_c$. The procedure generates a string of contributions of the type $(\partial)^{n_1}(\mathcal{M})^{n_2}(N_c)^{1-n_3}$, where n_1, n_2, and n_3 are non-negative integers. Order the expansion by first collecting all contributions with a given value of $n\equiv\frac{1}{2}n_1+n_2+n_3-1$, and only then sum over n. In this bookkeeping the first term in Eq. (11.9) corresponds to $n_1=2$, $n_2=0$, $n_3=0$, i.e., to $n=0$; the second and third terms are also of order $n=0$. The formula (11.9) is exact in the sense that it correctly describes all contributions to the effective Lagrangian at leading order, $n=0$. In particular, at leading order, the η' field enters the effective Lagrangian in the same manner as the ordinary Goldstone bosons, except for the mass term generated by the topological susceptibility of gluodynamics. For a discussion of the higher-order contributions to the effective Lagrangian, we refer the reader to [38].

For quarks of equal mass m, the Lagrangian (11.9) describes a set of N_f^2-1 degenerate "pions" and an η' with a different mass. Setting $\theta=0$ and taking m real, we obtain

$$M_\pi^2=\frac{2\Sigma m}{F^2} , \qquad M_{\eta'}^2=\frac{2}{F^2}\{N_f\tau+\Sigma m\} . \quad (11.10)$$

In the chiral limit, the spectrum thus contains the standard set of Goldstone bosons plus a particle whose mass is determined by the topological susceptibility. If N_c is sent to infinity, the η' also becomes massless, because the ratio τ/F^2 is of order $1/N_c$. Note that the first term in the curly brackets of Eq. (11.10) is small compared to the second as far as powers of N_c are concerned, while, comparing powers of the quark mass, the situation is reversed. In the bookkeeping introduced above, however, the two terms count as quantities of the same order: $M_{\eta'}^2$ is a small quantity of order $n=1$, irrespective of the relative magnitude of τ and Σm.

We now turn to the properties of the partition function. Suppose first that the box is large compared to the meson Compton wavelengths. The fluctuations in $U(x)$ then freeze, and the partition function reduces to the contribution from the ground state:

$$Z(\theta)=e^{-V\epsilon(\theta)} . \quad (11.11)$$

The vacuum-energy density is given by the minimum of the effective action:

$$\epsilon=\mathop{\rm Min}_U\left\{-\Sigma\, {\rm Re}\, {\rm tr}(\mathcal{M}U^\dagger)+\frac{\tau}{2}(\phi-\theta)^2\right\} \quad [U\in U(N_f)] . \quad (11.12)$$

If we again consider quarks of equal, positive mass m, the minimum occurs when U is a multiple of the unit matrix, $U=\exp(-i\phi/N_f)$, irrespective of the value of the vacuum angle. The formula (11.12) then reduces to

$$\epsilon={\rm Min}_\phi\left\{-\Sigma m N_f \cos\left[\frac{\phi}{N_f}\right]+\frac{\tau}{2}(\phi-\theta)^2\right\} . \quad (11.13)$$

For $\theta=0$, both terms in this expression are minimized by $\phi=0$. If θ is small, the minimum occurs at a small value of ϕ and we may therefore approximate the cosine by the first two terms of its Taylor series. In this approximation the minimum occurs at

$$\phi=\frac{N_f\tau\theta}{N_f\tau+\Sigma m}+O(\theta^3) . \quad (11.14)$$

In the chiral limit, where the susceptibility term dominates, the η' field sits at $\phi=\theta$. On the other hand, in the region $\Sigma m \gg N_f\tau$ where the η' becomes degenerate with the Goldstone bosons, ϕ vanishes and the ground state is described by $U=\mathbb{1}$. For the vacuum-energy density at small θ, Eq. (11.14) leads to

$$\epsilon=-N_f\Sigma m+\frac{1}{2}\theta^2\frac{\tau\Sigma m}{N_f\tau+\Sigma m}+O(\theta^4) . \quad (11.15)$$

The term quadratic in θ determines the topological susceptibility at $\theta=0$:

$$\frac{\langle \nu^2\rangle}{V}=\int dx\,\langle \omega(x)\omega(0)\rangle=\frac{\tau\Sigma m}{N_f\tau+\Sigma m} . \quad (11.16)$$

This shows what happens with the distribution of the winding number if N_c is sent to infinity. For $N_c=3$ and a small quark mass, we are in the region $\Sigma m \ll N_f\tau$ such that the mean-square winding number is given by $\langle \nu^2\rangle=V\Sigma m/N_f$, in agreement with what we found in Secs. IV and IX. Keeping the quark mass fixed and sending N_c to infinity, τ tends to a finite limit while Σm grows without bounds. We therefore wind up in the region $\Sigma m \gg \tau N_f$, where the mean-square winding number is given by $\langle \nu^2\rangle=V\tau$, as if we were dealing with the purely gluonic theory.

The above discussion concerns the behavior of the theory for boxes which are large compared to the Compton wavelengths of the lightest excitations, where the fluctuations in the winding number are Gaussian. Let us now turn to the region of small masses, where the finite-size effects generated by the box do become important. As discussed in detail in Sec. VIII for a finite number of colors, the fluctuations in the meson field $U(x)$ then show up. For $M_\pi L \ll 1$, $M_{\eta'} L \ll 1$, the nonzero modes contribute a term which represents the free energy of a massless Bose gas. Since this contribution does not depend on $\mathcal{M}$ or θ, it only affects the overall normalization constant. The fluctuations in the zero modes, on the other hand, are described by a collective integral over the group $\mathrm{U}(N_f)$:

$$Z(\theta)=B\int_{\mathrm{U}(N_f)} d\mu(U)\exp\left\{V\Sigma\,\mathrm{Re}\,\mathrm{tr}(\mathcal{M}U^\dagger)-\frac{V\tau}{2}(\phi-\theta)^2\right\} \tag{11.17}$$

[compare Eq. (8.10)]. The normalization constant B contains an extra factor from the η' degree of freedom:

$$B=\left[\frac{\pi F^2 V^{1/2} e^{\beta_0}}{N_f}\right]^{1/2} A\ , \tag{11.18}$$

where A is given in Eq. (8.11).

In the representation (11.17), the vacuum angle exclusively occurs in the susceptibility term. The Fourier integration with respect to θ is therefore Gaussian and can be done explicitly, with the result

$$Z_\nu=C_\nu\int_{\mathrm{U}(N_f)} d\mu(U)(\det U)^\nu e^{V\Sigma\,\mathrm{Re}\,\mathrm{tr}(\mathcal{M}U^\dagger)}\ . \tag{11.19}$$

Remarkably, the additional light meson which arises in the large-N_c limit merely affects the overall factor C_ν:

$$C_\nu=\frac{1}{\sqrt{2\pi V\tau}}e^{-\nu^2/2V\tau}B\ . \tag{11.20}$$

In other words, at a given winding number, the dependence of the partition function on the quark-mass matrix is determined by the same group integral as for finite N_c [compare Eq. (9.1)]. Accordingly, the spectrum of the Dirac operator at small eigenvalues also remains the same. In particular, the sum rules (9.21) thus also hold in the large-N_c limit. Note, however, that the scale of the eigenvalue distribution is set by the quark condensate which grows with N_c. Expressed in terms of the mass scale of the bound states, say, in terms of M_ρ, the spacing of the eigenvalues shrinks if the number of colors is increased. At $N_c=\infty$, the spectrum of the Dirac operator is continuous even if the volume is finite.

The occurrence of an extra light meson does show up in the distribution of the winding number. According to Eq. (11.20), the probability to encounter field configurations with large winding numbers acquires an additional suppression if N_c becomes large, through a Gaussian factor of width $\langle \nu^2\rangle_G=V\tau$. As discussed in detail in Sec. IX, the group integral also tends to zero for large values of ν, but this only happens if ν^2 reaches values of order $V\Sigma\mathcal{M}$ [recall that, for $N_f=1$, the group integral is given by the Bessel function $I_\nu(V\Sigma m)$]. For small quark masses and a small value of N_c, we are in the region $V\Sigma\mathcal{M} \ll V\tau$, such that the group integral falls off before the Gaussian factor connected with the topological susceptibility starts deviating from 1. As N_c grows, the width of the Gaussian remains the same, but the group integral falls off more slowly with ν because $V\Sigma\mathcal{M}$ is proportional to N_c. In the large-N_c limit, the susceptibility term wins, cutting the winding-number distribution off at values of order $\nu^2\sim V\tau$. The quark degrees of freedom strongly affect the properties of the partition function in the symmetry-restoration region $V\Sigma\mathcal{M}\lesssim 1$, but outside this region, the probability for finding field configurations of large winding number is the same as in gluodynamics.

If the fermions are in the *adjoint* representation of the gauge group, the fermion loops are of the same order in N_c as the gluon loops. In particular, the graphs responsible for the U(1) anomaly are then not suppressed in the large-N_c limit. For adjoint fermions the chiral-symmetry group is therefore not enlarged if N_c is sent to infinity and there is no reason for the analogue of η' to become massless. The analysis given in Sec. X does therefore not require any modifications in the large-N_c limit.

XII. SUMMARY AND CONCLUSION

The analysis described in the present paper leads to a rather detailed picture for the distribution of the winding number in QCD as well as for the spectrum of the Dirac operator at small eigenvalues. The main results are the following [39]. To simplify the discussion, we restrict ourselves to quarks of equal, positive mass m. [The significance of the phase of m is discussed in Sec. VII and the properties of the partition function for a quark-mass matrix of general form are analyzed in Sec. VIII.]

(i) We use the standard representation of the partition function in terms of an Euclidean path integral over gluon and quark fields on a torus. The periodicity condition for the gauge field implies that the winding number is an integer multiple of $1/N_c$. In QCD, however, only gauge-field configurations of integer winding number occur, because quarks transform according to the fundamental representation of the gauge group—the corresponding antiperiodicity condition cannot be met for gauge fields of fractional winding number. The partition function of QCD thus involves an integral over all gluon-field configurations of given integer winding number ν, followed by a sum over ν weighted with the factor $e^{i\nu\theta}$. [Although this definition of the partition function is natural and convenient, it is not mandatory. There are several alternative partition functions, i.e., alternative states, which, for quarks of nonzero mass, all lead to the same physics at infinite volume, but differ at finite volume. The issue is discussed in detail in Secs. II and X.]

(ii) Both the distribution of winding number and the spectrum of the Dirac operator at small eigenvalues are related to the quark condensate at infinite volume (massless quarks, $\theta=0$):

$$\Sigma=-\lim_{m\to 0}\lim_{V\to\infty}\langle \bar{u}u\rangle\ . \tag{12.1}$$

At finite volume the properties of the partition function are controlled by the parameter $x=V\Sigma m$. If x is large, the box does not significantly affect the properties of the system, but if x is small, the finite-size effects generated by the box strongly distort the behavior. In particular, if the chiral limit is taken at finite volume, one winds up at $x=0$; the properties of the partition function in that limit are drastically different from the physical situation which corresponds to quarks of nonzero mass at infinite volume, $x=\infty$.

(iii) In the presence of a single quark flavor, the mass spectrum contains a gap and the correlation functions rapidly decrease with distance, while for $N_f \geq 2$ Goldstone modes occur, generating long-range correlations if the quark masses are small. The qualitative difference originates in the fact that, for $N_f=1$, the quark condensate does not break the symmetry of the Lagrangian, while, for two or more flavors, Σ represents the order parameter of a spontaneously broken symmetry. Note that, for $N_f \geq 2$, the order of limits in Eq. (12.1) is essential: In the chiral limit at finite volume, the expectation value $\langle \bar{u}u \rangle$ vanishes, because spontaneous symmetry breakdown can only occur if the volume is infinite. For $N_f=1$, on the other hand, an interchange of these limits lead to the same result.

Remarkably, despite the pronounced flavor dependence seen in the physical spectrum, the distributions of the winding number and small Dirac eigenvalues are essentially independent of the number of flavors. In this respect the occurrence of a condensate manifests itself in almost the same manner, irrespective of whether or not this condensate belongs to an asymmetric ground state.

(iv) The results for the winding-number distribution are the following. If the parameter $V\Sigma m$ is large, the distribution is Gaussian with mean square $\langle \nu \rangle^2 = V\Sigma m / N_f$ [for unequal masses, the ratio m/N_f is replaced by the reduced mass; see Eq. (9.8)]. In the opposite limit, $V\Sigma m \ll 1$, nonzero winding numbers are strongly suppressed and the partition function is dominated by topologically trivial gauge-field configurations. The probability for encountering a configuration with winding number ν is given by $(V\Sigma m)^{|\nu|N_f} \mathcal{N}_\nu / \mathcal{N}_0$ [an explicit expression for the normalization factor $\mathcal{N}_\nu$ is given in Eq. (9.28)]. As is well known, the suppression originates in the occurrence of fermionic zero modes. On gauge-field configurations of winding number ν, the Dirac operator admits $|\nu|$ zero modes; for N_f quark flavors of mass m, the Dirac determinant is therefore proportional to $m^{|\nu|N_f}$. The above result shows, however, that this suppression factor is accompanied by an enhancement factor proportional to the power $V^{|\nu|N_f}$ of the volume. Nontrivial topologies are suppressed only if the product $V\Sigma m$ is a small number. In other words, the suppression is a finite-size effect of the same nature as the symmetry-restoration phenomena associated with the absence of spontaneous symmetry breakdown at finite volume. In the physical situation, where the parameter $V\Sigma m$ is large, nontrivial topologies are not suppressed—as mentioned above, the mean-square winding number is then large, of order $V\Sigma m$, and field configurations of very different winding numbers become equally probable. The fact that most of these configurations are accompanied by a large number of fermionic zero modes does not matter because the number of zero modes *per unit volume* tends to zero if V tends to infinity.

(v) Quite generally, global notions such as winding number are irrelevant if $V\Sigma m$ is large. One may restrict the path integral to topologically trivial field configurations, $\nu=0$; if $V\Sigma m \gg 1$, the result coincides with the full partition function at $\theta=0$, except for an irrelevant normalization factor. What is essential for the anomaly to generate a mass gap in the singlet channel is that, locally, the winding-number density fluctuates in accordance with the path-integral representation. For a detailed discussion of the U(1) problem, we refer the reader to Secs. V and XI, where we analyze the winding-number distribution in the large-N_c limit.

(vi) The relation of Banks and Casher [4] shows that, at infinite volume, the magnitude of the condensate is determined by the density of small eigenvalues of the Dirac operator, $\lambda \ll \Lambda_{QCD}$: The number of levels contained in the interval $\Delta\lambda$ per unit volume is equal to $\Sigma\Delta\lambda/\pi$. At finite volume, the spectrum is discrete. For gluon-field configurations of winding number ν, there are $|\nu|$ zero modes. The nonzero eigenvalues λ_n, on the other hand, depend on the particular gluon field under consideration. For a single, massless flavor, we find that the nonzero eigenvalues fluctuate around the values $\lambda_n = \pm\xi_n / V\Sigma$, where $\xi_1, \xi_2, \ldots$ are the zeros of the Bessel function $J_\nu(\xi)$. In particular, the lowest few eigenvalues are of order $1/V\Sigma$. Their distribution is sensitive to the winding number, the levels being pushed up if $|\nu|$ grows. We establish a set of sum rules which relate inverse moments of the eigenvalue distribution to the quark condensate and which become exact in the infinite-volume limit. Consider, e.g., the sum over all positive eigenvalues of $1/\lambda_n^2$. We show that, if the volume is large, this sum is dominated by the contributions from small eigenvalues, $\lambda \ll \Lambda_{QCD}$, and is determined by the value of the condensate, the number of flavors, and the winding number: $\Sigma'_n 1/\lambda_n^2 = \frac{1}{4}(V\Sigma)^2(N_f+|\nu|)^{-1}$. Outside the region of very small eigenvalues, the result of Banks and Casher also holds at finite volume: In the range $1/V\Sigma \ll \lambda_n \ll \Lambda_{QCD}$, the number of levels per unit volume is independent of the size of the box.

(vii) The extension of our analysis to fermions which transform according to the adjoint rather than the fundamental representation of the gauge group (supersymmetric Yang-Mills theory, for example) requires a few modifications which are discussed in Sec. X. Since the adjoint representation is real, the spectrum of the Dirac operator consists of pairs of degenerate levels related by charge conjugation. The extra symmetry gives rise to additional conserved currents. For N_f massless Dirac flavors in the adjoint representation, e.g., the Lagrangian is symmetric under the group SU($2N_f$). In contrast with QCD, the occurrence of a fermionic condensate now generates Goldstone bosons even if there is only a single Dirac flavor. Also, adjoint fermions can "live" on gluon-field configurations with a fractional winding number. If the path integral is extended over all gauge fields on a torus, it is periodic in θ only on an extended periodicity interval, $\theta \to \theta + 2\pi N_c$ (for details, see Sec. X). Incidentally, fractional winding numbers also play a role in pure Yang-Mills theory. As discussed in detail in Sect. XI, the large-N_c counting rules of gluodynamics are consistent with periodicity only if the path integral extends over *all* gauge-field configurations, including fractional winding numbers.

The entire analysis of the present paper is based on the assumption that a fermion condensate is formed, breaking chiral symmetry if $N_f > 1$. The most interesting question in this context is *why* such a condensate arises or, equivalently, why the fermionic level spectrum is much more dense at small eigenvalues ($\lambda_n \propto 1/V$) than it is the case for free particles ($\lambda_n \propto 1/L$). During the last ten years, this question has received considerable attention in the literature [40,41,3], which offers several qualitatively different attempts at identifying the relevant collective variables in the Euclidean path integral. In particular, a model which pictures the ground state as a liquid of instantons and anti-instantons has been developed in some detail [41]. In this model the localized zero modes, which occur for isolated (anti-)instantons, repel, forming a band of small eigenvalues of order $1/V$, as required. For the model to work, it is essential that the fermion determinant induce short-range correlations which shield the topological charges of the building blocks—since the topological susceptibility is small, of order Σm, the winding number must average out within a distance of order $1/M_{\eta'}$. It is also important that the model resemble a fluid rather than a crystal in the sense that long-range correlations should not occur. The Schwinger model provides a soluble example for which the correlation properties of the vacuum fields are qualitatively very similar [42].

ACKNOWLEDGMENTS

We are indebted to H. Banerjee, A. Zhitnitsky, P. Minkowski, H. B. Nielsen, G. Schierholz, E. V. Shuryak, T. Trappenberg, U.-J. Wiese, C. Wiesendanger, and A. Wipf for very useful discussions. This work was supported in part by Schweizerischer Nationalfonds.

APPENDIX A: WINDING NUMBER ON A TORUS

In this appendix we derive an explicit formula which expresses the winding number of a gauge-field configuration in terms of the transition function $\Omega_a(x)$. Since it is essential here that we are representing the field by means of a single patch, let us first check that, on a torus, any gauge field can be described in this manner.

The argument goes by induction in the number of dimensions. On a four-dimensional regions with the topology of a cube, the ordinary differential equation $(\partial_4+i\mathbf{G}_4)U(x)=0$ always admits solutions. Hence we can pick a gauge where $\mathbf{G}_4$ vanishes, such that the transition functions become independent of x^4. The problem therefore boils down to the question of whether or not several patches may be needed on three-dimensional cubes, which in turn reduces to the corresponding problem in two dimensions. One more step leads to a one-dimensional gauge field which can always be gauged away. Hence pathwork is not needed on manifolds with the topology of the cube, in any number of dimensions.

So we describe the gauge field by means of a single function $\mathbf{G}_\mu(x)$ defined on the cube $0\le x^\mu \le L_\mu$. The values on opposite sides of the cube are related by the periodicity condition (2.3). Denote the transition function which connects the side $x^\mu=0$ with $x^\mu=L_\mu$ by $\Omega_\mu(x)$ [in the notation of Sec. II, $\Omega_1\equiv\Omega_{a_1}$ with $a_1=(L_1,0,0,0)$]. The winding number can be expressed in terms of the one-form $G(x)=dx^\mu \mathbf{G}_\mu(x)$ as

$$\nu=\frac{1}{32\pi^2}\int dx^4 G^a_{\mu\nu}\tilde{G}^{\,a}_{\mu\nu}=\frac{1}{8\pi^2}\int \mathrm{tr}\{(dG+iG^2)^2\}\ . \quad (A1)$$

The integrand is a total derivative,

$$\nu=\frac{1}{8\pi^2}\int d\ \mathrm{tr}\left\{G\,dG+\frac{2i}{3}G^3\right\}\ , \quad (A2)$$

which contains four terms $d=d_1+\cdots+d_4$. Consider the first one, where the integration over x^1 yields

$$\nu_1=\frac{e_1}{8\pi^2}\int_{x^1=0}\mathrm{tr}\left\{\hat{G}\,d\hat{G}+\frac{2i}{3}\hat{G}^3-G\,dG-\frac{2i}{3}G^2\right\}\ . \quad (A3)$$

Here $\hat{G}=G(x+a_1)$ is the shifted field, the vector a_1 pointing in the direction of the first axis, with $|a_1|=L_1$. To keep track of signs, it is convenient to retain the Grassmann basis vectors $e_1,\ldots,e_4$ attached to the various differential forms. The integrand in Eq. (A3) is a three-form; the term e_1 in front of the integral is the Grassmann element, which remains from the integration over d_1.

Now $G(x+a_\mu)$ and $G(x)$ are related by the transition function $\Omega_\mu(x)$,

$$G(x+a_\mu)=\Omega_\mu(x)\{G(x)+\omega_\mu(x)\}\Omega_\mu^{-1}(x)\ , \quad (A4)$$

where ω_μ is the one-form

$$\omega_\mu=i\Omega_\mu^{-1}d\Omega_\mu\ . \quad (A5)$$

Inserting this in Eq. (A3), the terms which are quadratic or cubic in G cancel and we obtain

$$\nu_1=\frac{e_1}{8\pi^2}\int_{x^1=0}\mathrm{tr}\left\{dG\,\omega_1-iG\omega_1^2-\frac{i}{3}\omega_1^3\right\}\ . \quad (A6)$$

Using the property $d\omega_\mu=i\omega_\mu^2$, the terms linear in the gauge field can be written as a total derivative:

$$\nu_1=\frac{e_1}{8\pi^2}\int_{x^1=0}\left\{d\ \mathrm{tr}(G\omega_1)-\frac{i}{3}\omega_1^3\right\}\ . \quad (A7)$$

The derivative involves three terms $d=d_2+d_3+d_4$. The first one, e.g., gives

$$\nu_{12}=\frac{e_1e_2}{8\pi^2}\int_{x^1=x^2=0}\mathrm{tr}\{G(x+a_2)\omega_1(x+a_2)-G(x)\omega_1(x)\}\ . \quad (A8)$$

Using the periodicity condition once more, the integrand takes the form

$$\mathrm{tr}[G\{\Omega_2^{-1}\omega_1(x+a_2)\Omega_2-\omega_1\}+\omega_2\Omega_2^{-1}\omega_1(x+a_2)\Omega_2] , \tag{A9}$$

where we have suppressed those arguments which do not involve a shift. A similar two-dimensional integral over the plane $x^1=x^2=0$ also arises from ν_2. Using the consistency condition

$$\Omega_2(x+a_1)\Omega_1(x)=\Omega_1(x+a_2)\Omega_2(x) , \tag{A10}$$

one verifies that the corresponding one-forms obey

$$\Omega_2^{-1}\omega_1(x+a_2)\Omega_2-\omega_1=\Omega_1^{-1}\omega_2(x+a_1)\Omega_1-\omega_2 , \tag{A11}$$

where quantities without argument are to be evaluated at x. The relation (A11) implies that the term linear in G occurring in Eq. (A9) cancels against the corresponding contribution in ν_{21}. What remains are two- and three-dimensional integrals exclusively involving the transition function:

$$\nu=-\frac{1}{24\pi^2}\sum_{\mu}e_\mu\int_{x^\mu=0}\mathrm{tr}\{(\Omega_\mu^{-1}d\Omega_\mu)^3\}+\frac{1}{8\pi^2}\sum_{\mu,\nu}e_\mu e_\nu\int_{x^\mu=x^\nu=0}\mathrm{tr}\{d\Omega_\mu\Omega_\mu^{-1}\Omega_\nu^{-1}(x+a_\mu)d\Omega_\nu(x+a_\mu)\} . \tag{A12}$$

We add a remark concerning field configurations for which the consistency condition (A10) involves nontrivial Z factors such that ν may take fractional values. The above calculation goes through, unharmed, because the identity (A11) also holds in the presence of such factors—the formula (A12) remains valid. In fact, the gauge field transforms according to the adjoint representation of the gauge group. The periodicity condition (2.3) thus only involves the adjoint representation $D(\Omega_a)$ of the transition function, where the center of the group is mapped into 1. Accordingly, formula (A12) also holds if Ω_μ is replaced by the matrix $D(\Omega_\mu)$, except that the result is to be divided by $2N_c$ to account for the difference in the Casimir invariants of the adjoint and fundamental representations.

APPENDIX B: BEHAVIOR OF Z_ν AT SMALL MASS

Let us describe here how the result (9.27) is derived. Note first of all that the determinant (9.24) is an even function of ν so that we can assume $\nu>0$ without loss of generality. Substituting the leading term in the expansion (4.6a) for the Bessel function, we get

$$Z_\nu\sim\left[\frac{x}{2}\right]^{\nu N_f}\left[\prod_{k=0}^{N_f-1}(\nu+k)!\right]^{-1}D_\nu^{(N_f)} , \tag{B1}$$

where

$$D_\nu^{(N_f)}=\begin{vmatrix} R_{\nu+N_f-1}^{(N_f)} & R_{\nu+N_f-1}^{(N_f-1)} & \cdots & 1 \\ R_{\nu+N_f-2}^{(N_f)} & R_{\nu+N_f-2}^{(N_f-1)} & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ R_\nu^{(N_f)} & \cdots & \cdots & 1 \end{vmatrix} , \tag{B2}$$

with

$$R_\nu^{(N_f)}=\prod_{k=1}^{N_f-1}(\nu-k+1) . \tag{B3}$$

When going from Eq. (9.24) to Eq. (B1), we assumed at first that all indices of Bessel functions are positive, i.e., $\nu\geq N_f-1$. But the result (B1) is true for any ν. Let the index of a Bessel function become negative, the terms in the determinant involving that Bessel functions just do not contribute to the leading small-x behavior. The same is true for the determinant (B2) where the corresponding entry (B3) turns to zero in this case. The formula (B1) holds also for $N_f=1$ if we set

$$D_\nu^{(1)}=1 . \tag{B4}$$

Let us now calculate the determinant (B2). To this end we subtract from each row the next one and use the property

$$R_{\nu+1}^{(2)}-R_\nu^{(2)}=1 ,$$

$$R_{\nu+1}^{(N_f)}-R_\nu^{(N_f)}=(N_f-1)R_\nu^{(N_f-1)} \quad (N_f\geq 3) , \tag{B5}$$

$$D_0^{(N_f)}=(N_f-1)!D_\nu^{(N_f-1)} \quad (N_f\geq 2) . \tag{B6}$$

Solving this recurrent relation with the initial condition (B4), we obtain

$$D_\nu^{(N_f)}=\prod_{k=1}^{N_f-1}k! . \tag{B7}$$

Substituting it in Eq. (B1), we arrive at the result (9.27).

[1] For recent reviews, see S. Gottlieb, in *Lattice '90*, Proceedings of the International Symposium, Tallahassee, Florida, 1990, edited by U. M. Heller, A. D. Kennedy, and S. Sanielevici [Nucl. Phys. B (Proc. Suppl.) **20**, 247 (1991)]; F. Karsch, in *Quark-Gluon Plasma*, edited by R. C. Hwa (World Scientific, Singapore, 1991); D. Toussaint, "Progress in lattice QCD-91," University of Arizona Report No. AZPHTH/91-33 (unpublished).

[2] M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, Nucl. Phys. **B147**, 385 (1979); **B147**, 448 (1979); in *Vacu-*

um Structure and QCD Sum Rules, edited by M. A. Shifman (Elsevier, Amsterdam, 1992).
[3] E. V. Shuryak, *The QCD Vacuum, Hadrons and Superdense Matter* (World Scientific, Singapore, 1988).
[4] T. Banks and A. Casher, Nucl. Phys. **B169**, 103 (1980); E. Marinari, G. Parisi, and C. Rebbi, Phys. Rev. Lett. **47**, 1795 (1981).
[5] C. Callan, R. Dashen, and D. Gross, Phys. Lett. **63B**, 334 (1976); R. Jackiw and C. Rebbi, Phys. Rev. Lett. **37**, 172 (1976).
[6] P. Binétruy and M. K. Gaillard, Phys. Rev. D **32**, 931 (1985).
[7] J. Gasser and H. Leutwyler, Phys. Lett. B **184**, 83 (1987); **188**, 477 (1987); H. Leutwyler, in *Field Theory on the Lattice,* Proceedings of the International Symposium, Seillac, France, 1987, edited by A. Billoire *et al.* [Nucl. Phys. B (Proc. Suppl.) **4**, 248 (1988)].
[8] H. Neuberger, Phys. Rev. Lett. **60**, 889 (1988); in *Field Theory on the Lattice* [7], p. 501.
[9] P. Hasenfratz and H. Leutwyler, Nucl. Phys. **B343**, 241 (1990).
[10] J. Gasser and H. Leutwyler, Nucl. Phys. **B307**, 763 (1988).
[11] C. N. Yang and T. D. Lee, Phys. Rev. **87**, 404 (1952).
[12] R. J. Crewther, Phys. Lett. **70B**, 349 (1977).
[13] *Handbook of Mathematical Functions,* edited by M. Abramowitz and I. A. Stegun (Dover, New York, 1970).
[14] C. Jayewardena, Helv. Phys. Acta **61**, 636 (1988).
[15] R. J. Crewther, in *Field Theoretical Methods in Particle Physics,* edited by W. Rühl (Plenum, New York, 1980); G. A. Christos, Phys. Rep. **116**, 252 (1984).
[16] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields, Course of Theoretical Physics,* (Pergamon, Oxford, 1970), Vol. 2.
[17] J. Smit and J. Vink, Nucl. Phys. **B286**, 485 (1987); **B303**, 36 (1988); J. Vink, *ibid.* **B307**, 549 (1988); M. Gökeler *et al.*, *ibid.* **B334**, 527 (1990); S. Hands, J. Kogut, and E. Dagotto, *ibid.* **B333**, 551 (1990); I. Barbour *et al.*, *ibid.* **B368**, 390 (1992).
[18] D. Toussaint, in *Lattice '90* [1].
[19] J. Smit and J. Vink, Phys. Lett. B **194**, 433 (1987); J. Hoek, M. Teper, and J. Waterhouse, Nucl. Phys. **B288**, 589 (1987); A. S. Kronfeld *et al.*, *ibid.* **B292**, 330 (1987); A. Di Giacomo, in *QCD '90,* Proceedings of the International Workshop, Montpellier, France, 1990, edited by S. Narison [Nucl. Phys. B (Proc. Suppl.) **23B**, 191 (1991)].
[20] See, e.g., V. V. Khoze and A. V. Yung, Z. Phys. C **50**, 155 (1990).
[21] F. C. Hansen, Nucl. Phys. **B345**, 685 (1990); F. C. Hansen and H. Leutwyler, *ibid.* **B350**, 201 (1991).
[22] A. Hurwitz, Nachr. Ges. Wiss. Göttingen Math. Phys. Kl., 71 (1897).
[23] P. Di Vecchia and G. Veneziano, Nucl. Phys. **B171**, 253 (1980).
[24] H. Weyl, *The Classical Groups* (Princeton University Press, Princeton, NJ, 1946); S. Helgason, *Groups and Geometric Analysis* (Academic, New York, 1983).
[25] G. 't Hooft, Nucl. Phys. **B153**, 141 (1979); Acta Phys. Austriaca Suppl. **22**, 531 (1980).
[26] G. 't Hooft, Commun. Math. Phys. **81**, 267 (1981).
[27] P. Ramond, *Field Theory: A Modern Primer* (Benjamin/Cunnings, Reading, MA, 1981).
[28] A. I. Vainshtein and V. I. Zakharov, Pis'ma Zh. Eksp. Teor. Fiz. **35**, 258 (1982) [JETP Lett. **35**, 323 (1982)].
[29] E. Witten, Phys. Lett. **117B**, 324 (1982).
[30] E. Witten, Nucl. Phys. **B202**, 253 (1982).
[31] E. Cohen and C. Gomez, Phys. Rev. Lett. **52**, 237 (1984).
[32] A. R. Zhitnitsky, Nucl. Phys. **B340**, 56 (1990).
[33] V. A. Novikov *et al.*, Nucl. Phys. **B229**, 407 (1983); G. Rossi and G. Veneziano, Phys. Lett. **138B**, 195 (1984).
[34] G. 't Hooft, Nucl. Phys. **B72**, 461 (1974); **B75**, 461 (1974); G. Veneziano, *ibid.* **B117**, 519 (1974); E. Witten, *ibid.* **B160**, 57 (1980).
[35] E. Witten, Nucl. Phys. **B156**, 269 (1979); G. Veneziano, *ibid.* **B159**, 213 (1979); P. Di Vecchia, Phys. Lett. **85B**, 357 (1979); P. Nath and R. Arnowitt, Phys. Rev. D **23**, 473 (1981).
[36] S. Coleman and E. Witten, Phys. Rev. Lett. **45**, 100 (1980).
[37] C. Rosenzweig, J. Schechter, and T. Trahern, Phys. Rev. D **21**, 3388 (1980); E. Witten, Ann. Phys. (N.Y.) **128**, 363 (1980).
[38] J. Gasser and H. Leutwyler, Nucl. Phys. **B250**, 465 (1985).
[39] For a preliminary account, see H. Leutwyler, in *Effective Field Theories of the Standard Model,* Dobogokö, Hungary, 1991, edited by U.-G. Meissner (World Scientific, Singapore, 1992).
[40] C. G. Callan, R. Dashen, and D. J. Gross, Phys. Rev. D **17**, 2717 (1978); **19**, 1826 (1979); R. D. Carlitz and D. B. Creamer, Ann. Phys. (N.Y.) **118**, 429 (1979); E. Floratos and J. Stern, Phys. Lett. **119B**, 419 (1982); D. I. Diakonov and V. Yu. Petrov, *ibid.* **147B**, 351 (1984); Nucl. Phys. **B245**, 259 (1984); Yu. A. Simonov, Phys. Rev. D **43**, 3534 (1991).
[41] E. V. Shuryak and J. J. M. Verbaarschot, Nucl. Phys. **B341**, 1 (1990).
[42] A. V. Smilga, Phys. Rev. D **46**, 5598 (1992).
[43] E. Witten, Am. Phys. (N.Y.) **128**, 363 (1980); R. J. Crewther, Phys. Lett. **93B**, 75 (1980); Nucl. Phys. **B209**, 413 (1982).